www.ietdl.org

Published in IET Systems Biology
Received on 30th June 2010
Revised on 10th August 2010
doi: 10.1049/iet-syb.2010.0061

ISSN 1751-8849

Addressing parameter identifiability by model-based experimentation

A. Raue[1] *C. Kreutz*[1] *T. Maiwald*[2] *U. Klingmüller*[3] *J. Timmer*[4]

[1]*Physics Institute, University of Freiburg, 79104 Freiburg, Germany*
[2]*Department of Systems Biology, Harvard Medical School, 02115 Boston, MA, USA*
[3]*Division of Systems Biology of Signal Transduction, DKFZ-ZMBH Alliance, German Cancer Research Center, 69120 Heidelberg, Germany*
[4]*Physics Institute and Freiburg Institute for Advanced Studies, University of Freiburg, 79104 Freiburg, Germany*
E-mail: andreas.raue@fdm.uni-freiburg.de

Abstract: Mathematical description of biological processes such as gene regulatory networks or signalling pathways by dynamic models utilising ordinary differential equations faces challenges if the model parameters like rate constants are estimated from incomplete and noisy experimental data. Typically, biological networks are only partially observed. Only a fraction of the modelled molecular species is measurable directly. This can result in structurally non-identifiable model parameters. Furthermore, practical non-identifiability can arise from limited amount and quality of experimental data. In the challenge of growing model complexity on one side, and experimental limitations on the other side, both types of non-identifiability arise frequently in systems biological applications often prohibiting reliable prediction of system dynamics. On theoretical grounds this article summarises how and why both types of non-identifiability arise. It exemplifies pitfalls where models do not yield reliable predictions of system dynamics because of non-identifiabilities. Subsequently, several approaches for identifiability analysis proposed in the literature are discussed. The aim is to provide an overview of applicable methods for detecting parameter identifiability issues. Once non-identifiability is detected, it can be resolved either by experimental design, measuring additional data under suitable conditions; or by model reduction, tailoring the size of the model to the information content provided by the experimental data. Both strategies enhance model predictability and will be elucidated by an example application.

1 Introduction

'Those involved [in] estimating parameters from measurements would of course like to know whether they stand any chance of succeeding'. This statement in the foreword of the classics identifiability of parametric models [1] brings the whole issue of this article to the point. It is a topic becoming easily involved with sophisticated mathematical theory but, at the same time, it is of overwhelming practical importance. Last but not least, it should be possible to draw reliable conclusions from the mathematical framework utilised to describe experimental observations. This article tries to summarises the topic in a pragmatic and practical way, suitable to understand and overcome the complications implied by parameter non-identifiability issues.

Since its introduction in the late 1980s that was mostly driven by theoretical interests, parameter identifiability recently regained importance in the mathematical description of cellular processes that became known as systems biology. Bio-technological progress provided the possibility to decipher the human genome, yielding a multiplicity of genes representing specific functional components such as proteins or siRNA [2, 3]. However, the functionality observed in cellular processes is not arising from the components directly but from their dynamic interplay. Therefore the sequenced genome is rather an inventory of components than an explanation or 'manual' of functionality. The remaining task, putting together the known pieces to understand how functionality arises from the dynamic interplay of these components, is enormous. Detailed knowledge regarding this interplay is important to understand how failures in functionality arise and how resulting diseases such as cancer could be treated more efficiently. Since the underlying network of dynamic protein interactions is often only known in parts and because of its immense complexity, it is most suitable to describe the relevant processes by parametric models incorporating the biological knowledge available [4]. This statistical sound procedure promises reliable conclusions about the underlying biology, given that beforehand unknown model parameters such as rate constants of molecular interactions are reliably estimated from the available experimental data.

In the context of signalling networks, ordinary differential equations (ODE) are frequently used to investigate the dynamic properties of pathway components and their transient modifications. This assumes that diffusion is fast

compared to reaction rates and cell volume. The aim is to match the mathematical model with experimentally observed time-series data, to reconstruct and validate the network structure. An important step is the estimation of beforehand unknown model parameters determining the dynamical behaviour. Once an appropriate and reasonable model is established, it is usually desired to predict model dynamics such as trajectories of experimentally unobserved species concentrations or model behaviour under changed environmental conditions such as altered network structure or different external stimulation. Since the considered models are parametric, the outcome of these predictions intrinsically depends on the previously estimated model parameters and their identifiability.

Owing to technical limitations biological networks are frequently only partly accessible by experiments. As a consequence, not all molecular species incorporated in a model can be measured directly. For example, in the case of time-series data obtained by quantitative immunoblotting (western blotting), a widely used technique to examine proteins, the availability and specificity of antibodies limits the accessible species. Furthermore, measurement errors of experimental techniques are often substantial. Given a certain amount and quality of experimental data measured under specific experimental conditions, it is uncertain whether the model parameters can be estimated reliably. Frequently, experimental data are insufficient considering the size of the model, leading to non-identifiable parameters [5, 6]. Even identifiable parameters can only be determined within confidence intervals, which contain the true value of the parameter with a desired probability, see, for example [7]. The size of the confidence intervals depends on the amount and quality of experimental data. Confidence intervals of non-identifiable parameters are infinite indicating that they cannot be estimated from the available experimental data.

If model parameters are not well determined, the predicted model dynamics are also not. Consequently, a biological question that depends on inferring the dynamics might not be answerable. Therefore it is crucial to have detailed knowledge about parameter non-identifiabilities and uncertainties of the parameter estimates in general. Only then, false conclusions from model analysis can be prevented and the precision of model predictions can be assessed realistically.

1.1 Model for reaction networks

Reaction networks can be modelled using systems of ODE

$$\dot{\boldsymbol{x}}(t,\,\theta) = \boldsymbol{f}(\boldsymbol{x}(t,\,\theta),\,\boldsymbol{u}(t),\,\theta) \tag{1}$$

$$\boldsymbol{x}(0,\,\theta) = \boldsymbol{x}_0(\theta) \tag{2}$$

$$\boldsymbol{y}(t_i,\,\theta) = \boldsymbol{g}(\boldsymbol{x}(t_i,\,\theta),\,\theta) + \boldsymbol{\epsilon}_i \tag{3}$$

see, for example [8]. The internal model states $\boldsymbol{x}$ describe via the ODE system (1) the dynamics of n species such as concentrations of proteins in different phosphorylation states. Their dynamical behaviour may depend on an input function $\boldsymbol{u}(t)$ such as an external treatment with ligands and model parameters $\theta = \{\theta_1 \ldots \theta_l\}$ such as rate constants or initial concentrations. The internal model states are mapped to m model observables $\boldsymbol{y}$ via an observation function $\boldsymbol{g}$ in (3). The model observables are the quantities accessible by experiments measured at times t_i. They may depend on additional parameters such as scaling or offset parameters included in θ. Often, only a subset or combinations of the modelled species are accessible by experiments, meaning that $m < n$. The measurement noise distribution $\boldsymbol{\epsilon}_{ki}$ is assumed to be known, for example, being independently normally distributed: $\boldsymbol{\epsilon}_{ki} \sim \mathrm{N}(0,\,\sigma_{ki}^2)$. Each component of $\boldsymbol{f}$ in (1) is usually composed of a sum of reaction fluxes of several protein interactions, see, for example [9]. Equation (2) indicates that the initial conditions $\boldsymbol{x}(\boldsymbol{0})$ of the ODE system (1) might be dependent as well on the model parameters θ that need to be estimated from experimental data.

1.2 Parameter estimation

Commonly the model parameters θ are unknown and have to be estimated from experimental data. The agreement of experimental data $y_k^{\dagger}(t_i)$ with the observables predicted by the model $y_k(t_i,\,\theta)$ for parameters θ is measured by an objective function, commonly the weighted sum of squared residuals

$$\chi^2(\theta) = \sum_{k=1}^{m} \sum_{i=1}^{d_k} \frac{1}{\sigma_{ki}^2} (y_k^{\dagger}(t_i) - y_k(t_i,\,\theta))^2 \tag{4}$$

where d_k denotes the number of data-points for each observable $k = 1 \ldots m$, measured at time points t_i with $i = 1 \ldots d_k$. σ_{ki} are the corresponding measurement errors. The parameters can be estimated by

$$\hat{\theta} = \min_{\theta}[\chi^2(\theta)] \tag{5}$$

For normally distributed measurement noise this corresponds to maximum likelihood estimation (MLE) of θ because

$$\chi^2(\theta) = \mathrm{const} - 2\,\log(L(\theta)) \tag{6}$$

where

$$L(\theta) = \prod_{k=1}^{m} \prod_{i=1}^{d_k} \frac{1}{\sqrt{2\pi\sigma_{ki}^2}} \exp\left(-\frac{1}{2} \left(\frac{y_k^{\dagger}(t_i) - y_k(t_i,\,\theta)}{\sigma_{ki}} \right)^2 \right) \tag{7}$$

is the likelihood function. For a detailed discussion of MLE in the context of non-linear regression models, see for example [10].

Sometimes information about the distribution of the model parameters is available from independent experiments or from the literature. This information can be incorporated into the model by an empirical prior distribution that extends the likelihood. A normally distributed prior $\theta_j \sim \mathrm{N}(\theta_j^{\dagger},\,\sigma_{\theta_j}^2)$ for parameter θ_j yields a penalised objective function

$$\tilde{\chi}^2(\theta) = \chi^2(\theta) + \frac{1}{\sigma_{\theta_j}^2} (\theta_j^{\dagger} - \theta_j)^2 \tag{8}$$

extending (4). In the following, $\chi^2(\theta)$ as defined in (4) will be used as placeholder for the likelihood, see (6).

1.3 Confidence intervals

Given an appropriate model that sufficiently describes the available experimental data, the uncertainty of parameter

estimates in terms of confidence intervals can be assessed by analysing the shape of the likelihood.

A confidence interval $[\sigma_i^-, \sigma_i^+]$ of a parameter estimate $\hat{\theta}_i$ to a confidence level $1-\alpha$ signifies, that the true value θ_i^* is located within this interval with probability $1-\alpha$. In the following, large sample confidence intervals and small sample confidence intervals will be introduced. Sometimes these are referred to as asymptotic and finite sample confidence intervals, respectively.

1.3.1 Large sample confidence intervals: The most simple and handy description of the shape of the likelihood is its curvature evaluated at the estimated parameter values $\hat{\theta}$, for example, by the Hessian matrix

$$\boldsymbol{H} = \nabla^{\mathrm{T}} \nabla \chi^2(\theta)|_{\hat{\theta}} \tag{9}$$

Using the covariance matrix of the parameter estimates $\boldsymbol{C} = 2\boldsymbol{H}^{-1}$, confidence intervals known as standard intervals are given by

$$\sigma_i^{\pm} = \hat{\theta}_i \pm \sqrt{Q(\chi^2_{\mathrm{df}}, 1-\alpha)\,\boldsymbol{C}_{ii}} \tag{10}$$

where $Q(\chi^2_{\mathrm{df}}, 1-\alpha)$ is the $1-\alpha$ quantile of the χ^2_{df}-distribution with df, degrees of freedom [11]. The choice of df yields two different types of confidence intervals: $\mathrm{df}=1$ results in point-wise confidence intervals that hold individually for each parameter, $\mathrm{df}=\#\theta$ being the number of parameters results in simultaneous confidence intervals that hold jointly for all parameters.

Standard intervals allow a good approximation of the uncertainty of $\hat{\theta}_i$, if the amount of experimental data is large compared to $\#\theta$ and/or the measurement noise is small [10]. They are exact if the observables $\boldsymbol{y}(t_i, \theta)$ depend linearly on θ. However, even for the simplest reaction network described by a system of ODE, the observables depend non-linearly on θ. Furthermore, the amount and quality of time-series data for biological applications are often limited. Therefore standard intervals might not be appropriate. Fig. 1 illustrates the discrepancy of standard interval ellipsoid to the actual shape of the likelihood.

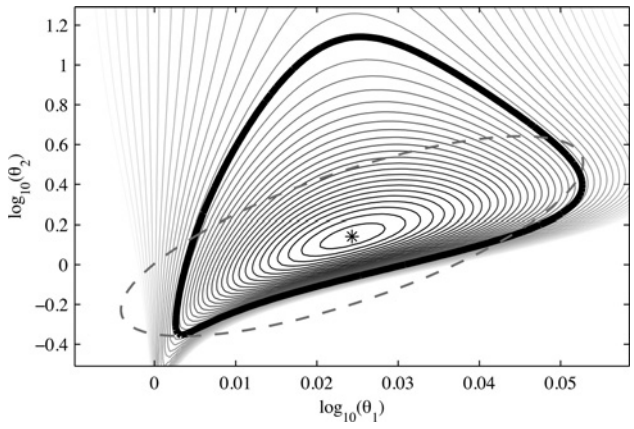


Fig. 1 *Contour plot of $\chi^2(\theta)$ on a $\log_{10}$-scale illustrating the discrepancy of standard intervals to the actual shape of the likelihood*

Contour lines shaded from black to white correspond to low, respectively, high values of $\chi^2(\theta)$. The asterisk indicates the minimum of χ^2 and the MLE for θ. The dashed ellipsoid indicates standard intervals and the thick contour line likelihood-based intervals

1.3.2 Small sample confidence intervals: Instead of using the local curvature of the likelihood, confidence intervals can be defined by a threshold Δ_α in the likelihood. This threshold defines a confidence region

$$\{\theta | \chi^2(\theta) - \chi^2(\hat{\theta}) < \Delta_\alpha\} \quad \text{with} \quad \Delta_\alpha = Q(\chi^2_{\mathrm{df}}, 1-\alpha) \tag{11}$$

whose borders represent likelihood-based confidence intervals [12]. The threshold Δ_α is the $1-\alpha$ quantile of the χ^2_{df}-distribution and represents with $\mathrm{df}=1$ and $\mathrm{df}=\#\theta$ point-wise, respectively, simultaneous confidence intervals to a confidence level of $1-\alpha$. The difference $\chi^2(\theta^*) - \chi^2(\hat{\theta})$ corresponds to the amount of over-fitting for the estimated parameters $\hat{\theta}$. As mentioned in [11], for non-linear models and small data samples the actual distribution of $\chi^2(\theta^*) - \chi^2(\hat{\theta})$ for the true parameters θ^* may differ from the χ^2_{df}-distribution. For instance, the distribution can be skewed, if the actual degrees of freedom df consumed by the non-linear model differ from the number of model parameters $\#\theta$. Since the deviation of the distribution is dependent on the model structure and the experimental conditions, the distribution of $\chi^2(\theta^*) - \chi^2(\hat{\theta})$ should always be verified by simulation studies. If deviations are observed, the threshold Δ_α should be adjusted according to the generated distribution. In any case, coverage properties of likelihood-based confidence intervals, that is, how precisely the confidence interval is matching the desired level of confidence, are considered superior to standard intervals for small samples [13]. For large samples both approaches are equivalent.

1.4 Identifiability

A parameter θ_i is structurally identifiable, if its estimate $\hat{\theta}_i$ is unique. It is practically identifiable, if the confidence interval of its estimate has finite size, that is, $\sigma_i^- > -\infty$ and $\sigma_i^+ < +\infty$ on a logarithmic scale that is usually used for rate constants or initial concentrations. A non-identifiable parameter indicates that it cannot be estimated from the experimental data and hence its confidence intervals are infinite. Structural non-identifiability is only related to the model structure and especially to the mapping $\boldsymbol{g}$ of internal model states to model observables in (3). It is independent of the accuracy of experimental data and was intensively discussed in the literature, for example, in [1]. In contrast, practical non-identifiability takes into account the amount and quality of experimental data that were used for parameter estimation. In the following, identifiability will be introduced from a data-based point of view rather than an algebraic one. In the last paragraph of this section, identifiability of model parameters will be linked to observability of model trajectories.

1.4.1 Structural non-identifiability: A structural non-identifiability arises from the model structure only and is independent of the amount and quality of experimental data. Consider a model defined by (1–3). Assuming ideal measurements, with arbitrarily many and perfectly chosen measurement time points t_i and the absence of measurements errors $\boldsymbol{\epsilon}_i = 0$, the crucial question is whether the model parameters θ are uniquely estimable from the model observables $\boldsymbol{y}(t_i, \theta)$.

The analytical solution of $\boldsymbol{y}(t_i, \theta)$ may contain an ambiguous parameterisation with respect to θ, arising from an insufficient mapping function $\boldsymbol{g}$ in (3). This ambiguity

doi: 10.1049/iet-syb.2010.0061

can be characterised as functional relations $\boldsymbol{h}(\theta_{\text{sub}}) = 0$ between a subset of parameters $\theta_{\text{sub}} \subset \theta$. Hence, the parameters θ_{sub} are structurally non-identifiable and can be varied according to the functional relations $\boldsymbol{h}$ without changing the observables $\boldsymbol{y}(t_i, \theta)$. In terms of an objective function such as $\chi^2(\theta)$ defined in (4), a structural non-identifiability manifests as iso-χ^2 manifold

$$\{\theta | \boldsymbol{h}(\theta_{\text{sub}}) = 0\} \quad \Rightarrow \quad \chi^2(\theta) = \text{const} \tag{12}$$

Consequently, the parameter estimates $\hat{\theta}_{\text{sub}}$ are not uniquely identified by measurements of $\boldsymbol{y}(t_i, \theta)$. Confidence intervals of a structurally non-identifiable parameter $\theta_i \in \theta_{\text{sub}}$ are infinite and hence θ_i cannot be estimated at all. For ODE models, a direct detection of an ambiguous parameterisation in the analytic form of $\boldsymbol{y}(t_i, \theta)$ is hampered, because (1) can only be solved explicitly in special cases.

For a two-dimensional parameter space, $\chi^2(\theta)$ can be visualised as landscape. Structural non-identifiability results in a perfectly flat valley, infinitely extended along the corresponding functional relation, as illustrated in Fig. 2*a*.

Since structural non-identifiability is independent of the accuracy of experimental data, it cannot be resolved by increasing the amount and quality of existing measurements. The only remedy is a qualitatively new measurement which alters the mapping function $\boldsymbol{g}$ in (3), usually by increasing the number of observed species. A parameter is structurally identifiable, if a unique minimum of $\chi^2(\theta)$ with respect to θ_i exists (see Figs. 2*b* and *c*).

1.4.2 Practical non-identifiability: A parameter that is structurally identifiable may still be practically non-identifiable. This can arise owing to insufficient amount and quality of experimental data or inappropriately chosen measurement time points. It manifests in a confidence interval that is infinite, although the likelihood has a unique minimum for this parameter. It is important to note, that the standard intervals of an estimate $\hat{\theta}_i$ are always finite because in this case $\boldsymbol{C}_{ii} > 0$, see (10). They are not suitable to characterise practical non-identifiability. Therefore [6] utilises likelihood-based confidence intervals. Here, a parameter is declared practically non-identifiable, if the likelihood-based confidence region (11) is infinitely extended in direction of θ_i. This means that the increase in $\chi^2(\theta)$ stays below the threshold Δ_α for a desired $1 - \alpha$ confidence level in direction of θ_i. Similar to structural non-identifiability, the flattening out of the likelihood can continue along a functional relation.

For a two-dimensional parameter space, a practical non-identifiability can be visualised as a relatively flat valley, which is infinitely extended. The height distance of the valley bottom to the lowest point at $\hat{\theta}$ never exceeds Δ_α, as illustrated in Fig. 2*b*.

The confidence interval of a practically non-identifiable parameter is not necessarily infinite on both sides. There can be a finite upper or lower bound of the confidence interval $[\sigma_i^-, \sigma_i^+]$, but either σ_i^- or σ_i^+ is infinite. For parameters such as rate constants or initial concentrations that are only positive, a log-transformation yields a more natural parameterisation. In this case, $\sigma_i^- = -\infty$ indicates that the true parameter θ_i^* could be arbitrarily small.

Improving the detection of characteristic dynamical behaviour by increasing the amount and quality of measured data and/or the choice of measurement time points t_i will ultimately remedy a practical non-identifiability, yielding finite likelihood-based confidence intervals, see Fig. 2*c*. If the amount of data is further increased, one finally arrives at the large sample case, where standard intervals are a reasonable measure of confidence. Inferring how to decrease confidence intervals most efficiently by additional experimental data is the subject of experimental design. How to propose new experiments that efficiently improve parameter identifiability and narrow confidence intervals will be addressed in Section 3: Consequences of non-identifiability.

1.4.3 Connection of identifiability and observability: The uncertainty of parameter estimates $\hat{\theta}$ directly translate

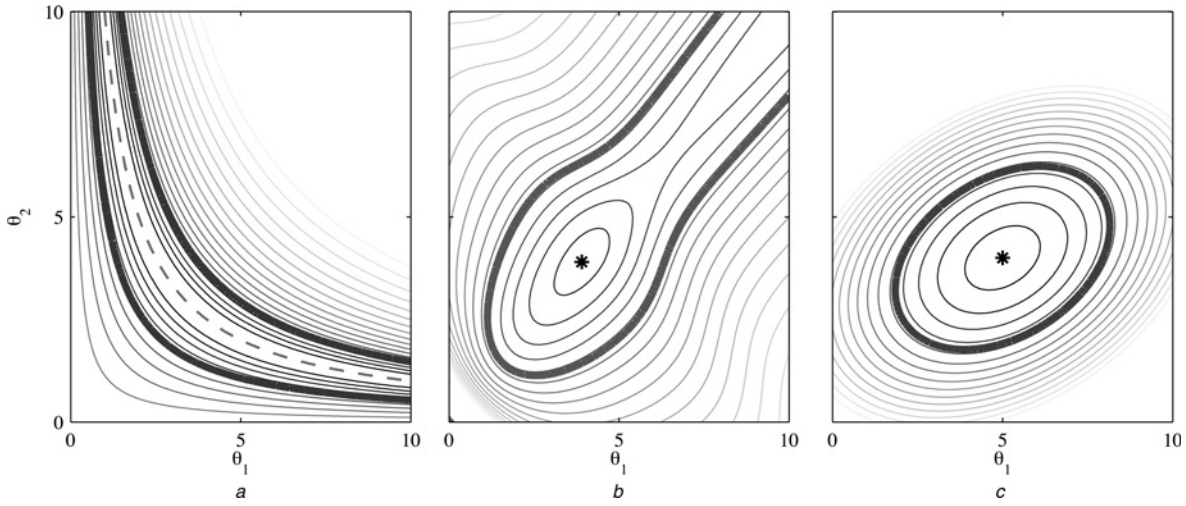


Fig. 2 *Illustrative contour plots of $\chi^2(\theta)$ for a two-dimensional parameter space*

a Structural non-identifiability along the functional relation $h(\theta_1, \theta_2) = \theta_1 \cdot \theta_2 - 10 = 0$ (dashed line). The likelihood-based confidence region is infinitely extended and there is no unique set of optimal parameters

b Practical non-identifiability along the functional relation $h(\theta_1, \theta_2) = \theta_1 - \theta_2 = 0$. The likelihood-based confidence region is infinitely extended along the functional relation for increasing values of θ_1 and θ_2. Lower-confidence bounds are well defined

c Both parameters are identifiable with finite-confidence intervals

Contour lines shaded from black to white correspond to low, respectively, high values of $\chi^2(\theta)$. Thick contour lines indicate likelihood-based confidence regions and asterisk the optimal parameters $\hat{\theta}$

to uncertainty of model trajectories. Non-identifiability describes the phenomena that parameters might not be determined. Consequently, non-observability indicates that model trajectories might not be determined owing to non-identifiability of model parameters.

Internal model states $x_j(t, \theta)$ that depend on structurally non-identifiable parameters θ_{sub} can be non-observable, whereas the model observables $\boldsymbol{y}(t_i, \theta)$ are per definition invariant. In contrast, practical non-identifiability does affect the model observables $\boldsymbol{y}(t_i, \theta)$. However, along practical non-identifiability, the model observables stay in agreement with the measurement accuracy of the experimental data because the likelihood stays below the threshold Δ_α. Nevertheless, some internal model states $\boldsymbol{x}(t, \theta)$ might be affected strongly by practical non-identifiability and hence might not be observable. In analogy to the type of identifiability, observability can be distinguished between structural or practical. Also, confidence intervals of parameter estimates translate to confidence intervals of model trajectories.

From a different point of view, identifiability can be restated in terms of observability. Additional internal model states x_{n+i} corresponding to parameters θ_i could be introduced by enlarging (1) with $\dot{x}_{n+i}(t) = 0$ and replacing every occurence of θ_i by x_{n+i} in (1–3). Consequently, showing the observability of x_{n+i} implies the identifiability of θ_i.

2 Methods

Various approaches to detect non-identifiability have been proposed. Early works focus on the analytical analysis of the model equations (1–3) whereas recently approaches utilise the growing amount of computing power for data-based analysis of parameter identifiability.

Approaches that analytically analyse the model equations are called a priori methods. They allow one to check for identifiability before experimental data are available. Consequently, these approaches allow one to check for structural identifiability only. Although structural identifiability might be ensured a priori, practical non-identifiability can cause severe problems when estimating model parameters from real experimental data [10]. Data-based approaches utilise the shape of the likelihood function to infer identifiability. This naturally involves experimental data and therefore allows statements about practical identifiability akin to confidence intervals.

Since the model equations are analysed analytically, a priori approaches have the advantage that conclusions about identifiability globally hold for the entire parameter space. Data-based approaches can, for reasons of computational expensiveness, not validate the entire parameter space [14]. Nevertheless [10] argues that local identifiability is usually sufficient in practice.

2.1 A priori approaches

In case (1–3) are only linearly depended on the model parameters and the dependency on the variables $\boldsymbol{x}$ is linear too, the model equations simplify to

$$\dot{\boldsymbol{x}}(t, \theta) = \boldsymbol{A}(\theta)\boldsymbol{x}(t, \theta) + \boldsymbol{B}(\theta)\boldsymbol{u}(t) \tag{13}$$

$$\boldsymbol{x}(0, \theta) = \boldsymbol{x}_0(\theta) \tag{14}$$

$$\boldsymbol{y}(t_i, \theta) = \boldsymbol{C}(\theta)\boldsymbol{x}(t_i, \theta) + \boldsymbol{\epsilon}_i \tag{15}$$

where matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ depend linear on the model parameters θ. For this restricted case some well-studied and efficient approaches exist.

The transfer function approach described, for example, in [10] assumes, besides linearity of the model equations, that matrices $\boldsymbol{B}$ and $\boldsymbol{C}$ are not parameter dependent. In this case, the analytical solution for the model observables is given by

$$\boldsymbol{y}(t_i, \theta) = \int_0^{t_i} \boldsymbol{C} \mathrm{e}^{\boldsymbol{A}(\theta)(t_i-\tau)} \boldsymbol{B}\boldsymbol{u}(\tau)\mathrm{d}\tau \tag{16}$$

Applying a Laplace transformation $\mathcal{L}_f(s) = \int_0^\infty \mathrm{e}^{-st} f(t)\mathrm{d}t$ with $s \in \mathbb{C}$, we obtain

$$\mathcal{L}_{\boldsymbol{y}}(s) = \underbrace{\boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A}(\theta))^{-1}\boldsymbol{B}}_{=\boldsymbol{\Phi}(s,\theta)} \mathcal{L}_{\boldsymbol{u}}(s) \tag{17}$$

The model parameters are structurally identifiable, if the transfer function $\boldsymbol{\Phi}(s, \theta)$, consisting of non-linear equations in powers of s and elements of $\boldsymbol{A}(\theta)$, have a unique solution for the parameters θ. However, solving the system of non-linear equations gets rapidly involved for larger models. A conceptually related approach, also for linear systems, uses the similarity transformation, see for example [15].

Another popular approach for linear systems utilises the Markov parameter matrix

$$\boldsymbol{M}(\theta) = \begin{bmatrix} \boldsymbol{C}(\theta)\boldsymbol{B}(\theta) \\ \boldsymbol{C}(\theta)\boldsymbol{A}(\theta)\boldsymbol{B}(\theta) \\ \vdots \\ \boldsymbol{C}(\theta)\boldsymbol{A}(\theta)^{2n-1}\boldsymbol{B}(\theta) \end{bmatrix} \tag{18}$$

where n is the number of internal model states, see for example [16]. The model parameters are structurally identifiable, if $\boldsymbol{M}(\theta)$ is of full rank. The approach is very efficient but only allows statements about local structural identifiability since $\boldsymbol{M}(\theta)$ is evaluated at a specific θ.

In biological applications the simplification to linear dynamical systems is often not indicated, for instance, if bi-molecular reactions or non-linear rate equations such as Michaelis–Menten kinetics are involved. A non-linear expansion of the transfer function approach was proposed by Lecourtier *et al.* [17]. An extension of the similarity transformation approach to non-linear systems using the local state isomorphism theorem was proposed by Vajda *et al.* [18].

A straightforward approach for non-linear dynamical systems is the power series expansion approach, introduced by Pohjanpalo [19]. Here time derivatives of the model observables are evaluated, yielding non-linear equations

$$\begin{aligned} \boldsymbol{y}(t, \theta) &= \boldsymbol{g}(\boldsymbol{x}(t, \theta), \theta) \\ \dot{\boldsymbol{y}}(t, \theta) &= \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{g}(\boldsymbol{x}(t, \theta), \theta) \\ \ddot{\boldsymbol{y}}(t, \theta) &= \frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{g}(\boldsymbol{x}(t, \theta), \theta) \\ &\vdots \end{aligned} \tag{19}$$

A model parameter θ_i is globally and structurally identifiable, if the system of equations provides an unique solution for this parameter. Similarly, this approach allows statement about the observability of the internal model states $\boldsymbol{x}$ and hence the

doi: 10.1049/iet-syb.2010.0061

identifiability of parameters accounting for the initial conditions $\boldsymbol{x}(0, \theta)$ via (2). Despite its conceptual clarity, the approach requires cumbersome calculation of high derivatives and large systems of non-linear equations to be solved. According to [20], the number of derivatives that need to be calculated is $(n + l)/m$. For the calculations, differential algebraic methods [21, 22] can be used. Furthermore, [23] proposed a probabilistic algorithm based on the power series expansion approach that address global structural identifiability and observability more efficiently and with high probability, also yielding the functional relations $\boldsymbol{h}$.

For all the so-far-mentioned a priori approaches, computational complexity is growing rapidly with increasing model size, compare [24, 25]. Therefore utilising computer algebra software is indispensable. In systems biological applications often models contain more than ten internal models states and more than twenty model parameters. The computational demand for a priori methods seems currently too large for these applications. Furthermore, a priori methods only regard structural identifiability. However, in applications it is essential to practically identify model parameters to ensure the estimation accuracy of the inferred dynamics in terms of confidence intervals.

2.2 Data-based approaches

The aim of data-based approaches is to detect non-identifiability by flatness of the likelihood function, using simulated or measured data. Owing to data dependency, the results of these methods cannot ensure global validity. Nevertheless, they allow statements about the region in parameter space specified by experimental data, which is the relevant region and usually sufficient for applications.

The most straightforward data-based approach for identifiability analysis is, from a perception of parameter estimation, akin to standard confidence intervals. Here, measures of curvature are computed, commonly using a quadratic approximation of $\chi^2(\theta)$ at the estimated optimum $\hat{\theta}$, for example, the Hessian (9) or the Fisher information matrix, see for example, in [26–29]. Flat directions in the likelihood that correspond to non-identifiability are characterised by zero eigenvalues of the utilised measure of curvature. Likewise, standard intervals are not suitable as confidence intervals, this approach faces serious problems in the assessment of identifiability for non-linear models in the small sample case: (i) The resulting linear parameter combinations might give a misleading impression of the potentially non-linear parameter relations $\boldsymbol{h}$ emerging from structural non-identifiability, see Fig. 3*a*. For reaction networks the parameter relations $\boldsymbol{h}$ are usually non-linear because of the dependency of the model observables on the model parameters through the ODE system. Furthermore, the threshold utilised to judge for 'zero' eigenvalues and 'flatness' of the likelihood, respectively, is not controllable in the non-linear case and yields results that are hardly interpretable [30]. (ii) Practical non-identifiability, as introduced above, cannot be detected because in that case the quadratic approximation does not fit the increasing but limited behaviour of $\chi^2(\theta)$. For practically non-identifiable parameters, this may result in incorrectly concluding structural non-identifiability, see Fig. 3*b*, or in incorrectly concluding identifiability, see Fig. 3*c*.

A remedy of the limitations imposed by the quadratic approximation, is to sample the likelihood more extensively. For a model with 20 parameters sampling the likelihood on a grid of size 100 would already require solving the ODE systems 10^{40} times, which is obviously not feasible. Markov chain Monte Carlo methods offer a more efficient sampling of the likelihood, but have convergence problems in the presence of non-identifiability because of the flat directions.

An alternative approach for identifiability analysis proposed by [6] utilises the profile likelihood

$$\chi^2_{\mathrm{PL}}(\theta_i) = \min_{\theta_{j\neq i}}[\chi^2(\theta)] \tag{20}$$

to infer flatness of the likelihood. Here, for each parameter θ_i individually, a section along the minimum of the objective

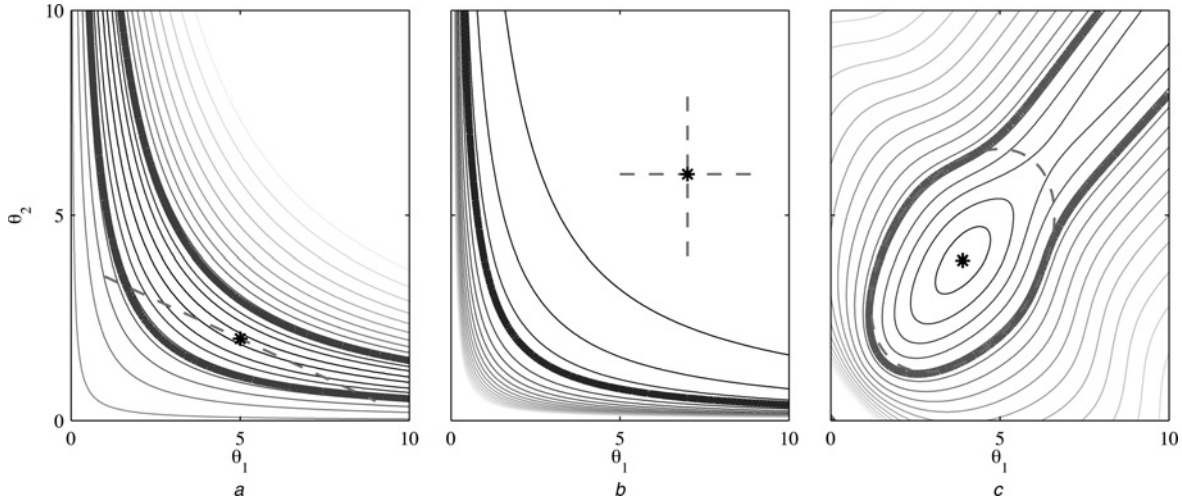


Fig. 3 *Problems when assessing identifiability from the quadratic approximation*

a Structural non-identifiability along the functional relation $h(\theta_1, \theta_2) = \theta_1 \cdot \theta_2 - 10 = 0$ is mistaken for a linear relationship, indicated by the dashed line
b Practical non-identifiability could be mistaken for a structural non-identifiability. Numerical estimation of the model parameters terminates when the likelihood becomes sufficiently flat, indicated by the asterisk point. The dashed lines correspond to 'flat' directions and 'zero' eigenvalues
c Two practical non-identifiable parameters are mistaken for identifiable, the dashed ellipsoid indicates standard intervals
Contour lines shaded from black to white correspond to low, respectively, high values of $\chi^2(\theta)$. Thick contour lines display likelihood-based confidence regions and asterisk correspond to the estimated parameters $\hat{\theta}$

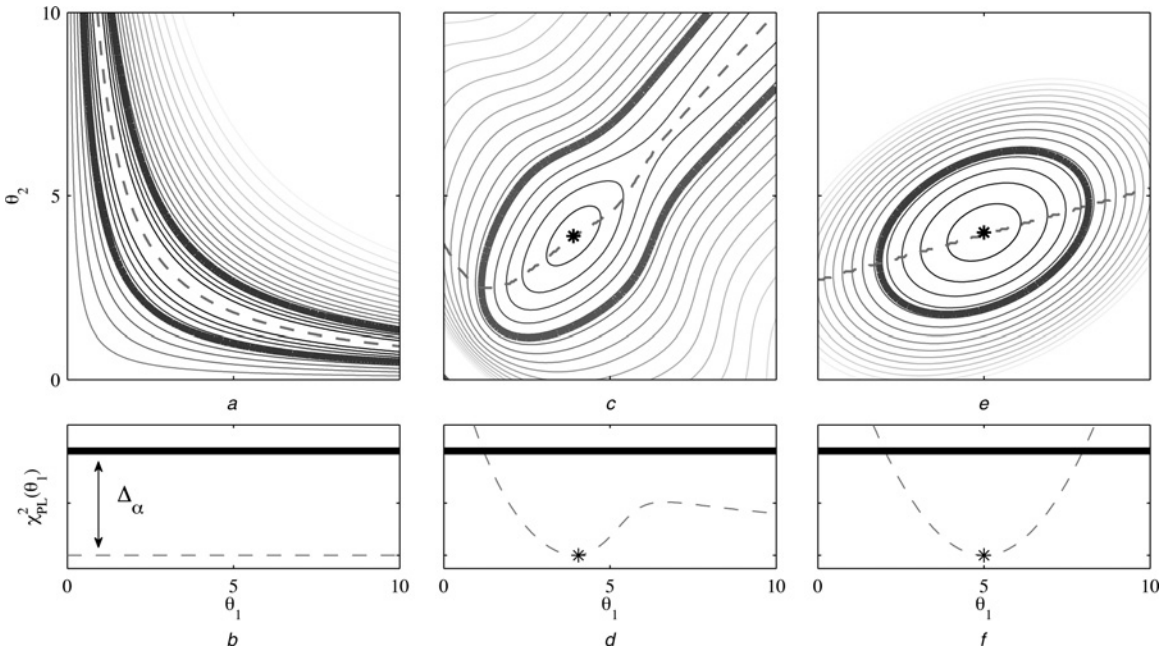


Fig. 4 *Assessing parameter identifiability of parameter θ_1 from the profile likelihood $\chi^2_{PL}(\theta_1)$*

a–b A structural non-identifiability along the functional relation $h(\theta_1, \theta_2) = \theta_1 \cdot \theta_2 - 10 = 0$ manifesting in a flat profile likelihood in panel *b*
c–d A practical non-identifiability manifests in a flatting out of the profile likelihood for $\theta_1 \to \infty$ in panel *d*
e The profile likelihood of an identifiable parameter θ_1
f The profile likelihood approaches a parabola shape indicating a good approximation by standard intervals
a, c, e The contour lines shaded from black to white correspond to low, respectively, high values of $\chi^2(\theta)$. Thick contour lines indicate likelihood-based confidence regions and asterisk correspond to the optimal parameters $\hat{\theta}$. Dashed lines indicate the trace of the profile likelihood for θ_1 in terms of parameter values
b, d, f The dashed lines indicate the profile likelihood χ^2_{PL} of parameter θ_1. The thick lines display the threshold Δ_α utilised to asses likelihood-based confidence regions for a confidence level α

function with respect to all of the other parameters $\theta_{j\neq i}$ is computed. It was originally proposed to calculate likelihood-based confidence intervals [31, 32]. It enables one to detect flatness of the likelihood function for arbitrary models and is especially useful in the small sample case. The idea of the approach is to explore the parameter space for each parameter in the direction of least increase in $\chi^2(\theta)$. Consequently, a structurally non-identifiable parameter manifests by a flat profile likelihood along the non-linear functional relations $\boldsymbol{h}(\theta_{\text{sub}}) = 0$, see Figs. 4*a* and *b*. In case of a practically non-identifiable parameter, the profile likelihood flattens out and does not exceed a given threshold defined by the desired confidence level, see Figs. 4*c* and *d*. In contrast, the profile likelihood of an identifiable parameter exceeds the threshold for both increasing and decreasing values of θ_i, see Figs. 4*e* and *f*. The points of passover represent likelihood-based confidence intervals [33]. The functional relations $\boldsymbol{h}$ connecting structurally non-identifiable parameters can be recovered from the profile likelihood by the change of the parameters $\theta_{j\neq i}$ while calculating the profile likelihood of the structurally non-identifiable parameter θ_i, see in [6].

An approach to infer the grouping of structural non-identifiable parameters owing to their functional relations $\boldsymbol{h}$ was introduced by [34]. It utilises the mean optimal transformation to detect non-linear relationships and is applied on samples obtained from minimising the likelihood.

3 Consequences of non-identifiability

Till now, identifiability was introduced and an overview about methods that allow for identifiability analysis was given. In applications, detecting a parameter non-identifiability is of course only the first step. Finally, the goal is to utilise an identifiable and hence observable model to predict and analyse dynamical systems behaviour with reasonable confidence. If a parameter non-identifiability is present, model predictions may not be reliable. In the remainder of this article, we want to discuss two frequently encountered scenarios, see Fig. 5 for an illustration.

Scenario 1: Depending on the scope of the investigation, the uncertainty introduced to model predictions owing to non-identifiability might affect a part of the model that is indispensable for the analysis. In this case, the only remedy to obtain reliable model predictions is to provide additional data measured under suitable experimental conditions by applying experimental design.

Scenario 2: If the part of the model affected by non-identifiability has none or only negligible effect on the part relevant for the investigation, the model may be reduced for the sake of simplifying the computational complexity.

In applications, it is important to know which of the two scenarios applies. To quantify the effect of a non-identifiable model parameter θ_i on the predicted dynamics, [6] proposed to plot the internal model trajectories $\boldsymbol{x}$ along the profile likelihood of this parameter.

3.1 Experimental design

When initial experiments are planned, usually the timescales of the participating molecular interactions are unknown. Therefore the measurement time points and target quantities are often not suitably chosen to unravel the system dynamics, resulting in non-identifiable model parameters. It would be valuable to suggest additional measurements that efficiently resolve parameter non-identifiability and narrow

doi: 10.1049/iet-syb.2010.0061

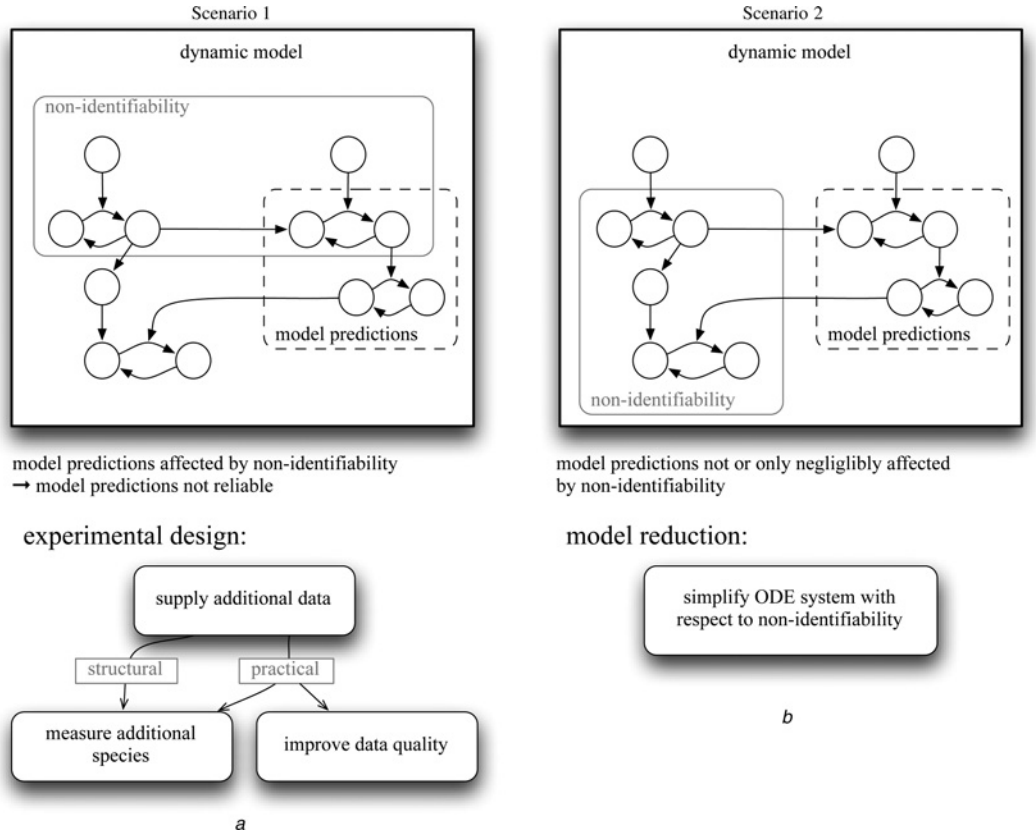


Fig. 5 *Two frequently encountered scenarios resulting from a non-identifiability*

a Uncertainty introduced to model predictions owing to non-identifiability affects a part of the model (both interactions and species framed by the non-identifiability box) that is indispensable for the model predictions (dashed box). The only remedy to obtain reliable model predictions is to provide additional data measured under suitable experimental conditions utilising experimental design

b Uncertainty imposed by non-identifiability affects not or only negligibly the part of the model that is relevant for the model predictions. The model may be reduced for the sake of simplifying the computational complexity

the confidence interval of a parameter θ_i affecting the dynamical behaviour of interest, consequently enhancing model predictability. For structural non-identifiability, it is not sufficient to improve the mere amount and quality of experimental data. Qualitatively new measurements that alter the observation function $\boldsymbol{g}$ in (3) are required, for example, measuring additional molecular species. For practical non-identifiability on the contrary, it is sufficient toincrease the amount and quality or the choice of measurement time points of existing measurements. Where feasible, qualitatively new measurements might still be more effective also in the case of practical non-identifiability.

Various methods exists that approach experimental design using the quadratic approximation of the likelihood, see for example, in [35] and applications in [36, 37]. In the case of non-linear models and small data samples these approaches suffer from similar problems as the corresponding methods for identifiability analysis. To overcome these problems, [6] proposed to investigate the set of trajectories along the profile likelihood of θ_i. This reveals spots where the uncertainty of θ_i has the largest impact on the model trajectories. Additional measurements at spots of largest variability promise to resolve both structural and practical non-identifiability and narrow confidence intervals most efficiently. Furthermore, the amplitude of variability of the trajectories at these spots allows one to assess the necessary measurement precision to provide adequate data. Methods for optimally finding the locations for additional measurements are also discussed in [38, 39]. The impact of new measurements can be evaluated by simulations before expensive experiments are carried out.

3.2 Model reduction

When building models, it is often unclear which features of a system are essential for determining its dynamical behaviour and which are dispensable. Consequently, detailedness of a model with respect to the scope of the investigation may be too large, frequently resulting in non-identifiabilities. Having ensured that model predictions are only negligibly affected by parameter non-identifiability, the ODE system (1) can be simplified. This reduces the computational complexity of further model analysis and yields a model tailored to the information content provided by the experimental data which remain predictive with respect to the scope of the investigation.

To quantify the effect of the uncertainty of a specific model parameter θ_i on the predicted dynamics, [6] proposed to plot the internal model trajectories $\boldsymbol{x}$ along the profile likelihood of this parameter. This allows one to evaluate which processes can be simplified without significantly altering the predicted dynamics. Whether the amount of uncertainty of the prediction is considered as significant depends on the application and the desired precision of the prediction.

Assume a reaction network modelled by mass action kinetics with a practically (or structurally) non-identifiable rate constant θ_i, to a point-wise confidence level by using $\mathrm{df} = 1$ in (11). Let us further assume that θ_i only negligibly

(or not at all) affects the model predictions relevant for the investigation. A (logarithmic) lower confidence bound of $\sigma_i^- = -\infty$ indicates that the corresponding reaction may be too slow to be detected on the timescale of the measurements. On the other hand, $\sigma_i^+ = +\infty$ indicates that the reaction may occurs too fast to be detected on the time scale of the measurements. The threshold Δ_α determining the confidence level corresponds to a likelihood ratio test, to a significance level of α, of the original model against a model reduced by the reaction corresponding to θ_i. Falling below this threshold, θ_i is practically (or structurally) non-identifiable and the likelihood ratio test indicates that it is not possible to dismiss the reduced model in favour of the original model. Hence, in both cases, given the information content provided by the experimental data, the model can be simplified by reducing the ODE system (1) without the model fit getting significantly worse. Note, that functional relations connecting θ_i with other parameters, as may be the case especially for the structural non-identifiable parameters, have to be taken into account.

3.3 Example application

Both strategies to resolve non-identifiabilities, experimental design and model reduction, will be demonstrated by a small example model and simulated data sets shown in Fig. 6, for a more realistic application see for example [40]. Briefly, an enzyme E is activated via two steps: E $\rightarrow$ *E and *E $\rightarrow$ **E. The first activation is dependent on a ligand L, modelled as an external input with constant magnitude of one. Once activated, the enzyme can catalyse substrate S to product P.

A typical situation in a systems biological application could be as followed: the product cannot be measured directly; the

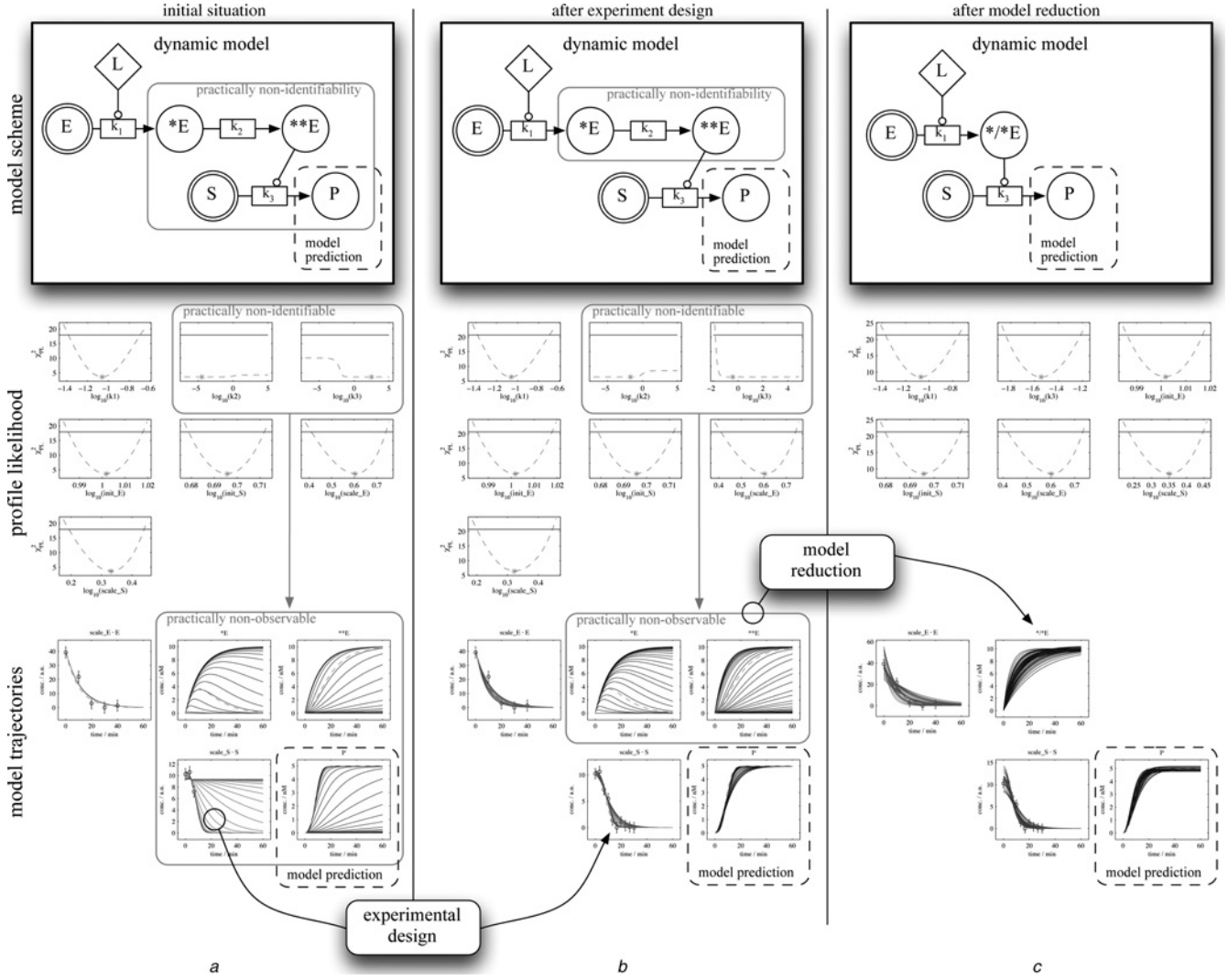


Fig. 6 *Illustrative example model and simulated data sets for demonstrating both strategies to resolve non-identifiabilities*

a After initial parameter estimation, the profile likelihood χ^2_{PL} of parameters k_2 and k_3 reveals a practical non-identifiability. Investigating the variability of the model trajectories for parameter values along $\chi^2_{PL}(k_3)$ reveals practical non-observability of species *E, **E, S and P. Consequently, the desired model prediction for P is not reliable. Additional measurements of S for $t > 10$ min are suggested to improve predictability

b Additional data allow for the identification of a lower-confidence bound for parameter k_3, critically determining the predicted dynamics of P. Parts of the practical non-identifiability of k_2 and k_3 remain, but only slightly affect the model prediction and hence allow for model reduction. The profile likelihood $\chi^2_{PL}(k_2)$ suggests that model-to-data agreement is compatible with $k_2 \rightarrow \infty$ suggesting unified modelling of *E and **E

c Simplified model is fully identifiable but also retains a reasonable prediction of P. Here, model trajectories indicate variability with respect to all model parameters simultaneously and hence evaluating confidence intervals of model trajectories

For interpretation of the 'profile likelihood' plots see Fig. 4 panels *b*, *d*, *f*. The asterisks indicate the parameter values after initial estimation. The horizontal line is the threshold Δ_α corresponding to simultaneous confidence intervals. The variability of the 'model trajectories' result from parameter uncertainties evaluated along the profile likelihood. The dashed trajectories indicate the true dynamics used to simulate the data set

doi: 10.1049/iet-syb.2010.0061

purpose of the investigation is the prediction of the dynamical behaviour of the product; measurable quantities are the inactive enzyme and the substrate concentrations only; both measurements have an unknown concentration scale but prior knowledge [see (8)] about absolute initial concentrations of E and S is available; internal model states represent species concentration in nM; measurements are sparse because of time and expensiveness. Furthermore, assume that no prior information about the reaction rate parameters, that is, the timescale of the system dynamics, is available. Consequently, initial experiments yield an 'explorative' data set that might not capture the underlying dynamics efficiently. Detailed information on the model structure and the data sets is available in the supplementary material.

The model is anticipated to reliably predict the products dynamics, not only for the current experimental conditions, but also for altered conditions, for example, altered ligand stimulation or a substance influencing one of the reaction rates. For this reason, the model should be fully identifiable. For the identifiability analysis the approach introduced in [6] was followed closely. The ODE system (1) was solved by CVODES [41], an efficient *C*-solver that allows us to compute the variational equations, also called sensitivity equations, simultaneously. For parameter estimation the MATLAB standard optimiser `lsqnonlin` was used. The example is available as repository for the PottersWheel multi-experiment fitting toolbox for MATLAB [42]. The calculation of the profile likelihood using this framework and for the given example take less than one minute per parameter on a 1.8 GHz dual core machine. For instructions how to reproduce the result obtained in the following, refer to the supplementary material.

The initial set-up is depicted in Fig. 6*a*. Initial parameter estimation yields a good model for data agreement with a value of the objective function $\chi^2 = 3.56$ for $8 + 2$ data points and 7 free parameters. Each prior information on the initial concentrations increases the expected value of χ^2 by one, see (8). As suggested in the section about likelihood-based confidence intervals, the actual distribution of $\chi^2(\theta^*) - \chi^2(\hat{\theta})$ was verified to be in accordance with the χ^2_7-distribution by a simulation study, refer to the supplementary material.

The key point of this article is that it is not sufficient to rely on the mere estimated parameter values and their corresponding prediction for the system dynamics. It is important to consider the uncertainties in the parameter estimation procedure: from measurement uncertainties, to parameter uncertainties and possibly non-identifiabilities, to uncertainties in the predicted model dynamics and possibly non-observabilities. Therefore after initial parameter estimation the profile likelihood of each parameter was evaluated, as indicated in Fig. 6*a*. Parameters k_2 and k_3 are revealed as practically non-identifiable, that is, their profile likelihood stays below the desired threshold for confidence intervals. Here, a 1σ simultaneous confidence level was used. A flat profile likelihood indicates that the model trajectories stay in agreement with the experimental data, defined by the measurement precision, for k_2 and k_3 taking extremely small and extremely large values. Here, the profile likelihood was evaluated between -5 and $+5$ on a $\log_{10}$-scale. The other parameters are identifiable and their confidence intervals are well defined. Investigating the variability of the model trajectories resulting from the practical non-identifiability of k_2 and k_3, as indicated in Fig. 6*a*, reveals practical non-observability of species *E, **E, S and P. Consequently, the desired model prediction for P is not reliable. A possible resolution to improve model predictability for P is to additionally measure S for $t > 10$ min.

The newly designed experiments, as indicated in Fig. 6*b* provide suitable data that allow for identification of a lower-confidence bound for parameter k_3. Parts of the practical non-identifiability of k_2 and k_3 still remain. Nevertheless, a repeated investigation of the variability of the model trajectories in response to practical non-identifiability of k_2 and k_3 confirms that the predicted dynamics for P is now only slightly affected. Consequently, the remaining non-identifiability may be removed from the model. The profile likelihood $\chi^2_{\mathrm{PL}}(k_2)$ suggests that model-to-data agreement is compatible with $k_2 \to \infty$ suggesting unified modelling of *E and **E.

The simplified model, as indicated in Fig. 6*c*, is now fully identifiable while retaining reasonable prediction of the dynamics of P. Here, model trajectories indicate variability with respect to all model parameters simultaneously and hence evaluating confidence intervals for the model prediction. The model fit in terms of the objective function increases from: $\chi^2 = 5.93$ for $15 + 2$ data-points and seven free parameters for the original model; to $\chi^2 = 7.39$ for $15 + 2$ data-points and six free parameters for the simplified model. The increase in the objective function, as already indicated by the profile likelihood, is insignificant in terms of a likelihood ratio test with $p = 0.2571$.

4 Summary

Parameter non-identifiability arise frequently in systems biological applications and are often insufficiently considered. We illustrated that parameter identifiability, both structural and practical, and confidence intervals of parameter estimates are a matter of flatness of the likelihood. For structural identifiability, it is critical as to which and how many of the modelled species can be measured directly. For practical identifiability, also the amount and quality of experimental data and the choice of the measurement time points play an important role. Both causes are inherent to biological applications where experiments are time consuming and expensive. Reliable parameter estimates are nevertheless critical before model predictions can be trusted. We discussed methods suitable for systems biological applications that allow us to assess and improve both structural and practical parameter identifiability in order to improve the reliability of model predictions.

Recently, it has been proposed that parameter non-identifiability manifested in ill-conditioned Hessian matrix with eigenvalues spanning orders of magnitudes, is an inherent property of the mechanistic ODE models utilised in systems biology [43]. However, the accuracy of the parameter estimates is related to both the model structure and the information provided by the experimental data. Non-identifiability predominantly arises owing to experimental restrictions. Without experimental limitations and if the models are reasonably parameterised in the common manner, non-identifiability issues can certainly be avoided by adequate experimental designs.

Using two frequently encountered scenarios, we illustrated the benefit of experimental design techniques in combination with models that are tailored to the information content provided by the experimental data, demonstrating that parameter identifiability is an issue that can be dealt with.

Thus, we would like to reformulate the introducing statement: those involved in estimating parameters from measurements do have a good chance of succeeding.

5 Acknowledgments

This work was supported by the German Federal Ministry of Education and Research [Virtual Liver, LungSys 0315415E, FRISYS 0313921]; the European Union [CancerSys EU-FP7 HEALTH-F4-2008-223188]; the Helmholtz Alliance on Systems Biology [SBCancer DKFZ I.2] and the Excellence Initiative of the German Federal and State Governments.

6 References

1 Walter, E.: 'Identifiability of parametric models' (Pergamon Press, 1987)
2 International Human Genome Sequencing Consortium: 'Finishing the euchromatic sequence of the human genome', *Nature*, 2004, **431**, (7011), pp. 931–945
3 Gerstein, M.B., Bruce, C., Rozowsky, J.S., *et al.*: 'What is a gene, post-encode? History and updated definition', *Genome Res.*, 2007, **17**, (6), pp. 669–681
4 Kitano, H.: 'A brief overview: systems biology', *Science*, 2002, **295**, (5560), pp. 1662–1664
5 Swameye, I., Müller, T.G., Timmer, J., Sandra, O., Klingmüller, U.: 'Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling', *Proc. Nat. Acad. Sci.*, 2003, **100**, (3), pp. 1028–1033
6 Raue, A., Kreutz, C., Maiwald, T., *et al.*: 'Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood', *Bioinformatics*, 2009, **25**, (15), pp. 1923–1929
7 Lehmann, E.L., Leo, E.: 'Theory of point estimation' (Wiley, 1983)
8 Wolkenhauer, O.: 'Systems biology' (Portland Press, 2008)
9 Heinrich, R., Schuster, S.: 'The regulation of cellular systems' (Chapman & Hall, 1996)
10 Seber, G.A.F., Wild, C.J.: 'Nonlinear regression' (Wiley, 2003)
11 Press, W.H., Teukolsky, S.L.A., Flannery, B.N.P., Vetterling, W.M.T.: 'Numerical recipes: FORTRAN' (Cambridge University Press, 1990)
12 Meeker, W.Q., Escobar, L.A.: 'Teaching about approximate confidence regions based on maximum likelihood estimation', *Am. Stat.*, 1995, **49**, (1), pp. 48–53
13 Neale, M.C., Miller, M.B.: 'The use of likelihood-based confidence intervals in genetic models', *Behav. Genet.*, 1997, **27**, (2), pp. 113–120
14 Audoly, S., D'Angio, L., Saccomani, M.P., Cobelli, C.: 'Global identifiability of linear compartmental models – a computer algebra algorithm', *IEEE Trans. Biomed. Eng.*, 1998, **45**, (1), pp. 36–47
15 Jacquez, J., Greif, P.: 'Numerical parameter identifiability and estimability: integrating identifiability, estimability, and optimal sampling design', *Math. Biosci.*, 1985, **77**, (1), pp. 201–228
16 Cobelli, C., DiStefano, J.J. III.: 'Parameter and structural identifiability concepts and ambiguities: a critical review and analysis', *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, 1980, **239**, (1), pp. 7–24
17 Lecourtier, Y., Lamnabhi-Lagarrigue, F., Walter, E.: 'Volterra and generating power series approaches to identifiability testing', in Walter, E. (Ed.): 'Identifiability of parametric models' (Pergamon Press, 1987)
18 Vajda, S., Godfrey, K.R., Rabitz, H.: 'Similarity transformation approach to identifiability analysis of nonlinear compartmental models', *Math. Biosci.*, 1989, **93**, pp. 217–248
19 Pohjanpalo, H.: 'System identifiability based on the power series expansion of the solution', *Math. Biosci.*, 1978, **41**, (1–2), pp. 21–33
20 Sadik, B.: 'A bound for the order of characteristic set elements of an ordinary prime differential ideal and some applications', *Appl. Algebra Eng.*, 2000, **10**, (3), pp. 251–268
21 Ljung, L., Glad, T.: 'On global identifiability for arbitrary model parametrizations', *Automatica*, 1994, **30**, pp. 265–265
22 Saccomani, M.P., Audoly, S., D'Angio, L.: 'Parameter identifiability of nonlinear systems: the role of initial conditions', *Automatica*, 2003, **39**, (4), pp. 619–632
23 Sedoglavic, A.: 'A probabilistic algorithm to test local algebraic observability in polynomial time', *J. Symbol. Comput.*, 2002, **33**, (5), pp. 735–755
24 Margaria, G., Riccomagno, E., Chappell, M.J., Wynn, H.P.: 'Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences', *Math. Biosci.*, 2001, **174**, (1), pp. 1–26
25 White, L.J., Evans, N.D., Lam, T., *et al.*: 'The structural identifiability and parameter estimation of a multispecies model for the transmission of mastitis in dairy cows', *Math. Biosci.*, 2001, **174**, (2), pp. 77–90
26 Vajda, S., Rabitz, H., Walter, E., Lecourtier, Y.: 'Qualitative and quantitative identifiability analysis of nonlinear chemical kinetic models', *Chem. Eng. Commun.*, 1989, **83**, (1), pp. 191–219
27 Yao, K.Z., Shaw, B.M., Kou, B., McAuley, K.B., Bacon, D.W.: 'Modeling ethylene/butene copolymerization with multi-site catalysts: parameter estimability and experimental design', *Polym. React. Eng.*, 2003, **11**, (3), pp. 563–588
28 Degenring, D., Froemel, C., Dikta, G., Takors, R.: 'Sensitivity analysis for the reduction of complex metabolism models', *J. Process Control*, 2004, **14**, (7), pp. 729–745
29 Jaqaman, K., Danuser, G.: 'Linking data to models: data regression', *Nat. Rev. Mol. Cell Biol.*, 2006, **7**, pp. 813–819
30 Quaiser, T., Monnigmann, M.: 'Systematic identifiability testing for unambiguous mechanistic modeling – application to JAK-STAT, MAP kinase, and NF-κB signaling pathway models', *BMC Syst. Biol.*, 2009, **3**, p. 50
31 Venzon, D.J., Moolgavkar, S.H.: 'A method for computing profile-likelihood-based confidence intervals', *Appl. Stat.*, 1988, **37**, (1), pp. 87–94
32 Murphy, S.A., van der Vaart, A.W.: 'On profile likelihood', *J. Am. Stat. Assoc.*, 2000, **95**, (450), pp. 449–485
33 Royston, P.: 'Profile likelihood for estimation and confidence intervals', *Stat. J.*, 2007, **7**, (3), pp. 376–387
34 Hengl, S., Kreutz, C., Timmer, J., Maiwald, T.: 'Data-based identifiability analysis of non-linear dynamical models', *Bioinformatics*, 2007, **23**, (19), pp. 2612–2618
35 Kreutz, C., Timmer, J.: 'Systems biology: experimental design', *FEBS J.*, 2009, **276**, (4), pp. 923–942
36 Schlatter, R., Conzelmann, H., Gilles, E., Sawodny, O., Sauter, T.: 'Analysis of an apoptotic core model focused on experimental design using artificial data', *IET Syst. Biol.*, 2009, **3**, (4), pp. 255–265
37 Bandara, S., Schlöder, J., Eils, R., Bock, H., Meyer, T.: 'Optimal experimental design for parameter estimation of a cell signaling model', *PLoS Comput. Biol.*, 2009, **5**, (11), pp. 1–12
38 Kutalik, Z., Cho, K., Wolkenhauer, O.: 'Optimal sampling time selection for parameter estimation in dynamic pathway modeling', *Biosystems*, 2004, **75**, (1–3), pp. 43–55
39 Balsa-Canto, E., Alonso, A., Banga, J.: 'Computational procedures for optimal experimental design in biological systems', *IET Syst. Biol.*, 2008, **2**, (4), pp. 163–172
40 Becker, V., Schilling, M., Bachmann, J., *et al.*: 'Covering a broad dynamic range: information processing at the erythropoietin receptor', *Science*, 2010, **328**, (5984), pp. 1404–1408
41 Hindmarsh, A.C., Brown, P.N., Grant, K.E., *et al.*: 'Sundials: suite of nonlinear and differential/algebraic equation solvers', *ACM Trans. Math. Softw.*, 2005, **31**, (3), pp. 363–396
42 Maiwald, T., Timmer, J.: 'Dynamical modeling and multi-experiment fitting with potterswheel', *Bioinformatics*, 2008, **24**, (18), pp. 2037–2043, available at http://www.potterwheel.de
43 Gutenkunst, R.N., Waterfall, J.J., Casey, F.P., Brown, K.S., Myers, C.R., Sethna, J.P.: 'Universally sloppy parameter sensitivities in systems biology models', *PLoS Comput. Biol.*, 2007, **3**, (10), pp. 1871–1878

doi: 10.1049/iet-syb.2010.0061

Supplementary Material: Addressing Parameter Identifiability by Model-Based Experimentation

A. Raue [1,*] , C. Kreutz [2] , T. Maiwald [3] , U. Klingmüller [4] , J. Timmer [5]

August 5, 2010

[1]Physics Institute, University of Freiburg, 79104 Freiburg, Germany, andreas.raue@fdm.uni-freiburg.de
[*]corresponding author
[2]Physics Institute, University of Freiburg, 79104 Freiburg, Germany, ckreutz@fdm.uni-freiburg.de
[3]Department of Systems Biology, Harvard Medical School, 02115 Boston, MA, USA, tmaiwald@googlemail.com
[4]Division of Systems Biology of Signal Transduction, DKFZ-ZMBH Alliance, German Cancer Research Center, 69120 Heidelberg, Germany, u.klingmueller@dkfz-heidelberg.de
[5]Physics Institute and Freiburg Institute for Advanced Studies, University of Freiburg, 79104 Freiburg, Germany, jeti@fdm.uni-freiburg.de

1 Model definition

Model parameters θ: {k1, k2, k3, init_E, init_S, scale_E, scale_S}

Input function $\vec{u}(t)$: $[\mathrm{L}] = 1$

ODE system $\dot{\vec{x}}(t, \theta)$, see main text (1):

$$\left(\begin{array}{rcl} [\dot{\mathrm{E}}] & = & -\mathrm{k1} \cdot [\mathrm{E}] \cdot [\mathrm{L}] \\ [^{*}\dot{\mathrm{E}}] & = & +\mathrm{k1} \cdot [\mathrm{E}] \cdot [\mathrm{L}] - \mathrm{k2} \cdot [^{*}\mathrm{E}] \\ [^{**}\dot{\mathrm{E}}] & = & +\mathrm{k2} \cdot [^{*}\mathrm{E}] \\ [\dot{\mathrm{S}}] & = & -\mathrm{k3} \cdot [^{**}\mathrm{E}] \cdot [\mathrm{S}] \\ [\dot{\mathrm{P}}] & = & +\mathrm{k3} \cdot [^{**}\mathrm{E}] \cdot [\mathrm{S}] \end{array}\right)$$

Initial conditions for ODE system $\vec{x}(0, \theta)$, see main text (2):

$$\begin{pmatrix} [\mathrm{E}](0) &=& \mathrm{init\_E} \\ [^{*}\mathrm{E}](0) &=& 0 \\ [^{**}\mathrm{E}](0) &=& 0 \\ [\mathrm{S}](0) &=& \mathrm{init\_S} \\ [\mathrm{P}](0) &=& 0 \end{pmatrix}$$

Model observables $\vec{y}(t_i, \theta)$, see main text (3):

$$\begin{pmatrix} y_1(t_i) &=& \mathrm{scale\_E} \cdot [\mathrm{E}](t_i) + \mathrm{N}(0,\, 4^2) \\ y_2(t_i) &=& \mathrm{scale\_S} \cdot [\mathrm{S}](t_i) + \mathrm{N}(0,\, 1^2) \end{pmatrix}$$

The measurement noise is assumed to be independent and normally distributed with known variance.

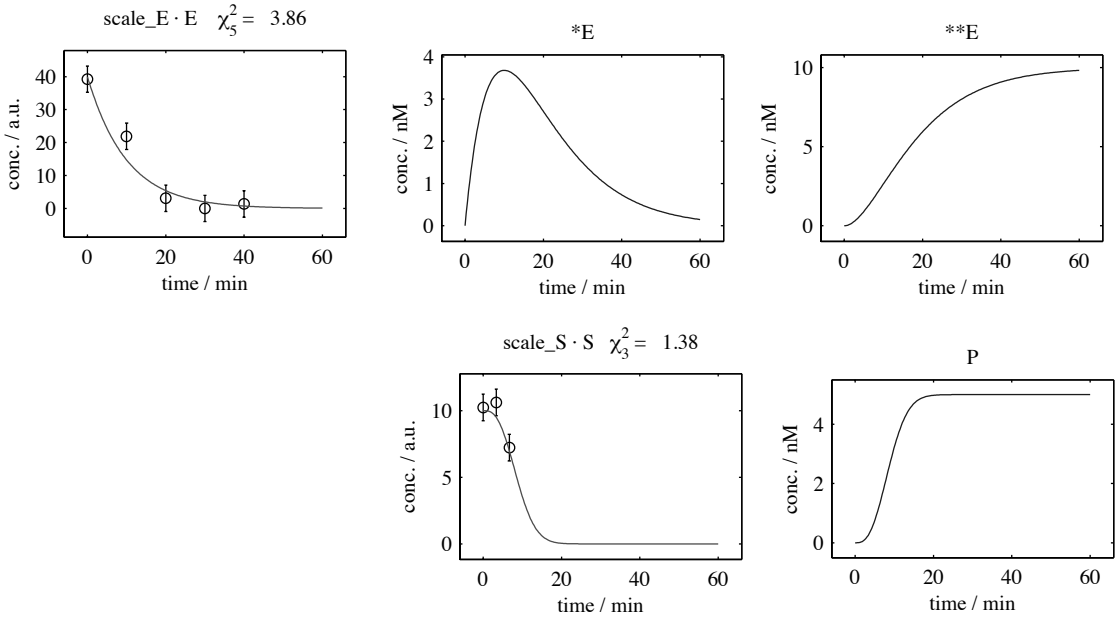
Prior information about model parameters:

$$\begin{aligned} \mathrm{init\_E} &\sim \mathrm{N}(10.04,\, 0.1^2) \\ \mathrm{init\_S} &\sim \mathrm{N}(4.96,\, 0.05^2) \end{aligned}$$

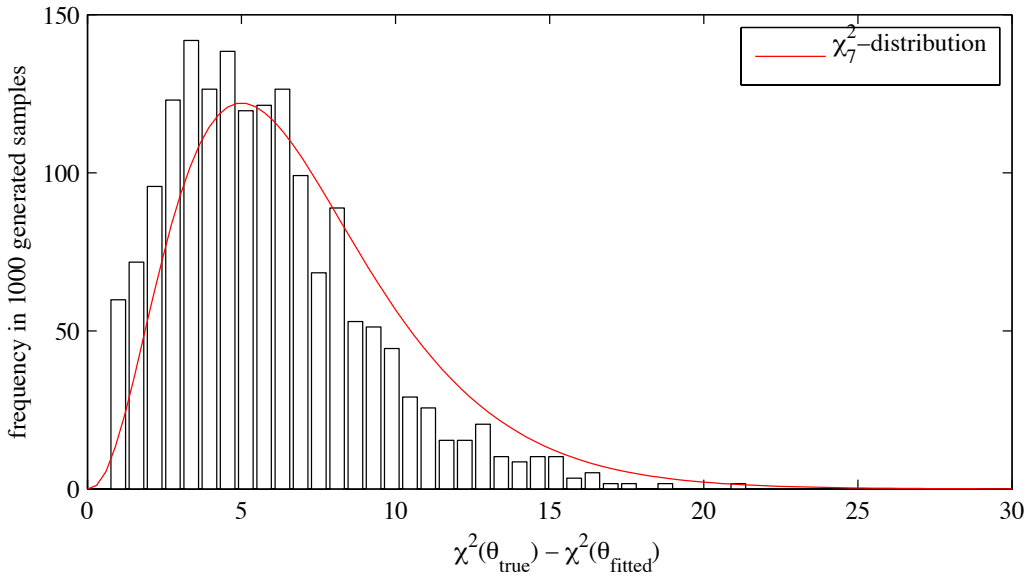
Initial data set for setup (a):

time / min	y_1 / a.u.	y_2 / a.u.
0	39.24	10.24
3.33	-	10.61
6.66	-	7.23
10	21.89	-
20	3.08	-
30	0.12	-
40	1.37	-

Plot of initial data set and true model trajectories:



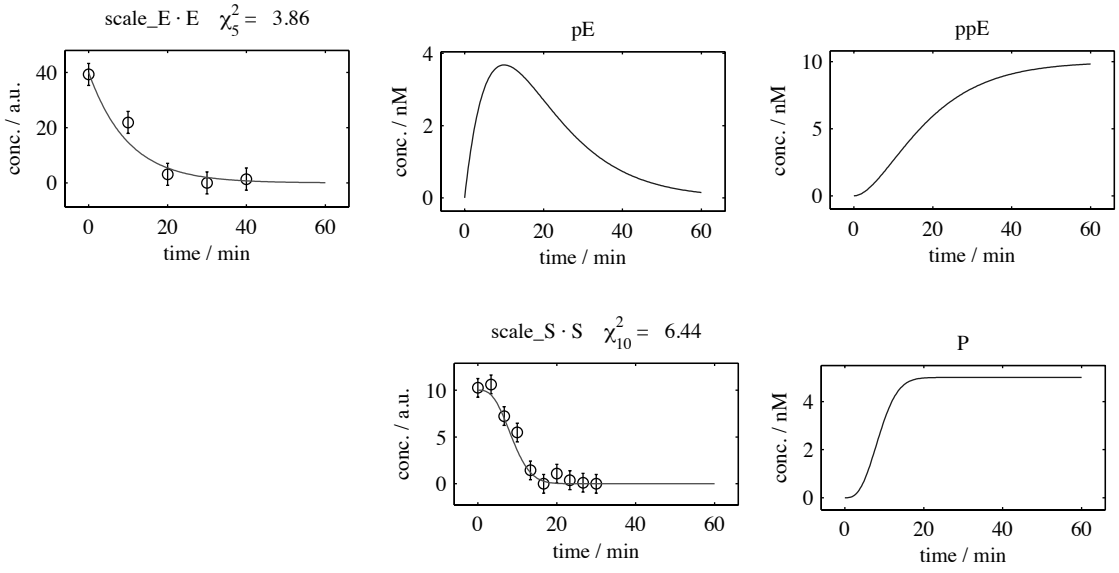
As suggested in the section about likelihood-based confidence intervals, the actual distribution of $\chi^2(\theta^*) - \chi^2(\hat{\theta})$ was verified to be in good agreement with the χ^2_{df}-distribution with degrees of freedom $df = \#\theta = 7$ by a simulation study using 1000 generated noise realizations:



Extended data set for setup (b) and (c):

time / min	y_1 / a.u.	y_2 / a.u.
0	39.24	10.24
3.33	-	10.61
6.66	-	7.23
10	21.89	5.47
13.33	-	1.43
16.66	-	0.53
20	3.08	1.08
23.33	-	0.38
26.66	-	0.11
30	0.12	0.08
40	1.37	-

Plot of extended data set and true model trajectories:



Simplified ODE system with $[^{*/*}\mathrm{E}] = [^{*}\mathrm{E}] + [^{**}\mathrm{E}]$:

$$\begin{pmatrix} [\dot{\mathrm{E}}] &=& -\mathrm{k1}' \cdot [\mathrm{E}] \cdot [\mathrm{L}] \\ [\dot{^{*/*}\mathrm{E}}] &=& +\mathrm{k1}' \cdot [\mathrm{E}] \cdot [\mathrm{L}] \\ [\dot{\mathrm{S}}] &=& -\mathrm{k3}' \cdot [^{*/*}\mathrm{E}] \cdot [\mathrm{S}] \\ [\dot{\mathrm{P}}] &=& +\mathrm{k3}' \cdot [^{*/*}\mathrm{E}] \cdot [\mathrm{S}] \end{pmatrix}$$

Parameter values on a $\log_{10}$-scale for setup (a) and (b):

name	θ^*	$\hat{\theta}_{(a)}$	$\sigma^-_{(a)}$	$\sigma^+_{(a)}$	$\hat{\theta}_{(b)}$	$\sigma^-_{(b)}$	$\sigma^+_{(b)}$
k1	-1	-1.062	-1.384	-0.6961	-1.061	-1.384	-0.695
k2	-1	-4.501	-Inf	+Inf	-1.719	-Inf	+Inf
k3	-1	+2.502	-Inf	+Inf	-0.431	-1.835	+Inf
init_E	+1	+1.002	+0.9851	+1.018	+1.002	+0.985	+1.018
init_S	+0.6990	+0.696	+0.6787	+0.7119	+0.696	+0.679	+0.712
scale_E	+0.6021	+0.604	+0.4095	+0.7413	+0.605	+0.410	+0.741
scale_S	+0.3010	+0.332	+0.1925	+0.4398	+0.322	+0.201	+0.446

Parameter values on a $\log_{10}$-scale for setup (c):

name	$\hat{\theta}_{(c)}$	$\sigma^-_{(c)}$	$\sigma^+_{(c)}$
k1'	-1.071	-1.372	-0.749
k3'	-1.545	-1.842	-1.214
init_E	+1.002	+0.986	+1.017
init_S	+0.696	+0.680	+0.711
scale_E	+0.603	+0.420	+0.734
scale_S	+0.351	+0.234	+0.448

2 PottersWheel instructions

The example application presented in the main text, section 3 "Consequences of non-identifiability", is available as repository for the PottersWheel multi-experiment fitting toolbox for MATLAB [1]. For instructions how to set up and use the toolbox see http://www.potterswheel.de. Please ensure that the CVODES solver [2] is setup correctly, follow the instruction given in `help pwInstallSundialsTB`. The model definition and data files necessary to reproduce the results for the presented example can be downloaded from http://web.me.com/andreas.raue/profile/software.html. Extract the archive provided there to a convenient location, start MATLAB and change the MATLAB working path to the newly created folder. In the following, all necessary function calls will be explained.

Initial situation For loading the initial experimental setup as depicted in Fig. 6a of the main text run the script `initial_setup.m` or load the repository `pwRepository_initial_setup.mat` using the PottersWheel user interface. The model trajectories and data sets can be plotted by the command `pwDraw`. To estimate the model parameters and calibrate the model trajectories to the data set execute the command `pwFit`. Alternatively, the repository `pwRepository_initial_setup_fitted.mat` with already estimated parameters values can be loaded.

To calculate the profile likelihood for the model parameters execute the script `ple_initial_setup.m`. The calculation takes about 24 ± 12 seconds per parameter on a 1.8 GHz dual core machine. The result are displayed on the screen and stored in a subfolder of the current directory labeled with a time code. For further background on the calculation of the profile likelihood, see [3]. To display the corresponding variability of the model trajectories for parameter values along the profile likelihood $\chi^2_{PL}(k_3)$ execute the command `pwPLETrajectories(3,3)`.

After experimental design For loading the experimental setup with the extended data set as depicted in Fig. 6b of the main text run the script `extended_data_setup.m` or load the repository

`pwRepository_extended_data_setup.mat` using the PottersWheel user interface. The model trajectories and data sets can be plotted by the command `pwDraw`. To estimate the model parameters and calibrate the model trajectories to the data set execute the command `pwFit`. Alternatively, the repository `pwRepository_extended_data_setup_fitted.mat` with already estimated parameters values can be loaded.

To calculate the profile likelihood for the model parameters execute the script `ple_extended_data_setup.m`. The calculation takes about 43 ± 33 seconds per parameter on a 1.8 GHz dual core machine. The result are displayed on the screen and stored in a subfolder of the current directory labeled with a time code. For further background on the calculation of the profile likelihood, see [3]. To display the corresponding variability of the model trajectories for parameter values along the profile likelihood $\chi^2_{PL}(k_3)$ execute the command `pwPLETrajectories(3,3)`.

After model reduction For loading the experimental setup with the extended data set as depicted in Fig. 6b of the main text run the script `simplified_model_setup.m` or load the repository `pwRepository_simplified_model_setup.mat` using the PottersWheel user interface. The model trajectories and data sets can be plotted by the command `pwDraw`. To estimate the model parameters and calibrate the model trajectories to the data set execute the command `pwFit`. Alternatively, the repository `pwRepository_simplified_model_setup_fitted.mat` with already estimated parameters values can be loaded.

To calculate the profile likelihood for the model parameters execute the script `ple_simplified_model_setup.m`. The calculation takes about 10 ± 2 seconds per parameter on a 1.8 GHz dual core machine. The result are displayed on the screen and stored in a subfolder of the current directory labeled with a time code. For further background on the calculation of the profile likelihood, see [3]. To display the variability of the model trajectories for parameter values along the profile likelihood of all model parameters execute the command `pwPLETrajectories(0,3)`.

References

[1] T Maiwald and J Timmer. Dynamical modeling and multi-experiment fitting with potterswheel. *Bioinformatics*, 24(18):2037–2043, 2008. http://www.potterwheel.de.

[2] AC Hindmarsh, PN Brown, KE Grant, SL Lee, R Serban, DE Shumaker, and CS Woodward. Sundials: Suite of nonlinear and differential/algebraic equation solvers. *ACM T Math Software*, 31(3):363–396, 2005.

[3] A Raue, C Kreutz, T Maiwald, J Bachmann, M Schilling, U Klingmuller, and J Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, 2009.