Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Systems Biology

On structural and practical identifiability

Franz-Georg Wieland[1,2,3], Adrian L. Hauber[1,2], Marcus Rosenblatt[1,2], Christian Tönsing[1,2,3] and Jens Timmer[1,2,3]

Abstract

We discuss issues of structural and practical identifiability of partially observed differential equations, which are often applied in systems biology. The development of mathematical methods to investigate structural nonidentifiability has a long tradition. Computationally efficient methods to detect and cure it have been developed recently. Practical nonidentifiability, on the other hand, has not been investigated at the same conceptually clear level. We argue that practical identifiability is more challenging than structural identifiability when it comes to modeling experimental data. We discuss that the classical approach based on the Fisher information matrix has severe shortcomings. As an alternative, we propose using the profile likelihood, which is a powerful approach to detect and resolve practical nonidentifiability.

Addresses

[1] Institute of Physics, University of Freiburg, Hermann-Herder-Str. 3, 79104, Freiburg, Germany

[2] Freiburg Center for Data Analysis and Modelling (FDM), University of Freiburg, Ernst-Zermelo-Str. 1, 79104, Freiburg, Germany

[3] Centre for Integrative Biological Signalling Studies (CIBSS), University of Freiburg, Schänzlestr. 18, 79104, Freiburg, Germany

Corresponding author: Timmer, Jens (jeti@fdm.uni-freiburg.de)

Current Opinion in Systems Biology 2021, **25**:60–69

This review comes from a themed issue on **Mathematical Modelling (2021)**

Edited by **Stacey D. Finley** and **Vassily Hatzimanikatis**

For complete overview of the section, please refer the article collection - Mathematical Modelling (2021)

Available online 22 March 2021

https://doi.org/10.1016/j.coisb.2021.03.005

Keywords

Identifiability, Structural identifiability, Practical identifiability, Profile likelihood, Fisher information matrix, Nonlinear dynamics, ODE models, Experimental design, Model reduction, Observability.

Introduction

Biological modeling and Box's statement

Traditional biological reasoning often is rather qualitative, descriptive, and static, which results, for example, in cell biology in so-called 'pathway cartoons'. Mathematical models based on differential equations can help to turn these into a quantitative, predictive, and dynamic understanding of the underlying system. Discussing modeling in general, in 1979, George E.P. Box coined his famous statement: "All models are wrong, but some are useful" [1]. Although the former part of the quote is intuitively clear, because every model necessarily poses a simplification of reality, the latter highlights the importance of assessing what constitutes a *useful* model.

Bad, good, and useful models

Three properties comprise a *useful* model. First, it has to capture the main effects of the question of interest, that is, describe the data with reasonable accuracy, and neglect the rest. Second, a *useful* model has to make experimentally falsifiable predictions to be testable. Models that exhibit these two properties are *good* models. Third, the model should enable to gather insights about the biological system. In a typical modeling process, one starts off with an initial model based on current biological knowledge. Usually, this model cannot explain the data and therefore is a *bad* model. On the basis of biological intuition and trial-and-error, one increases the model complexity until the data can be fitted. Often, this leads to an over-parameterized model that overfits the data. The parameters of such a model and, in turn, its predictions are not well determined, and it thus remains a *bad* model.

The path from such a *bad* model toward a *good* model is laborious: additional data need to be measured and integrated, the model complexity needs to be reduced and balanced to the available data, or a combination of both. This process needs to be iterated until a *good* model is found, which has well-determined parameters and predictions.

However, such a *good* model also needs to deliver biological insights in order to be *useful*. Only this third property turns a *good* model into a *useful* model. In this sense, the final goal of mathematical modeling in systems biology is not the model itself but to use the model to understand biology. One example of how a model can be used to gain biological insight, which would be unattainable by merely assessing the data by itself, was given by Becker et al. [2].

Parameter identifiability

The concept of identifiability is strongly linked to the transition from *bad* models to *good* models. Identifiability analysis is necessary to create *good* models that can describe the data and have well-determined parameters and predictions. It is especially important when modeling biological systems because the limited amount and quality of the experimental data with large measurement noise in only partially observed systems often lead to *bad* models during the modeling process. Concerning identifiability, one distinguishes between structural identifiability dealing with inherently indeterminable parameters because of the model structure itself, and practical identifiability, dealing with insufficiently informative measurements to determine the parameters with adequate precision.

Partially observed dynamical systems

A biological system is translated into ordinary differential equations (ODEs)

$$\dot{x} = f(x, p, u), \tag{1}$$

comprising n model states $x(t)$, unknown parameters p to be estimated from time-resolved experimental data, and external stimuli $u(t)$. As data is often recorded on a relative scale, scaling and offset parameters for background corrections need to be estimated in parallel. Furthermore, in typical applications, not all components of a cell-biological system can be measured, for example, because of the limited availability or restricted capability of antibodies to discriminate between unphosphorylated, that is, inactive, and phosphorylated, that is, active, proteins. Thus, an observation function $g(\cdot)$ is required, which maps the internal states x to the observations:

$$y = g(x, p, t). \tag{2}$$

Typically, the dimension m of y is smaller than the dimension n of x. We are therefore dealing with parameter estimation in partially observed systems. Moreover, in systems biology, these ODE models are typically stiff, nonlinear, sparse, and nonautonomous, and the discrete-time observations are noisy.

Parameter estimation is usually performed based on the weighted residual sum of squares, the negative log-likelihood assuming Gaussian errors

$$\chi^2_{\mathrm{res}}(p) = \sum_{k=1}^{m} \sum_{l=1}^{d_k} \left(\frac{y^D_{kl} - g_k(p, t_l)}{\sigma^D_{kl}} \right)^2, \tag{3}$$

to determine the agreement of experimental data with the model trajectories, where y^D_{kl} and σ^D_{kl} represent d_k data points and measurement errors at time points t_l for each observable.

A common point estimate for the best parameter vector is the maximum likelihood estimator

$$\hat{p} = \arg\min\left[\chi^2_{\mathrm{res}}(p)\right]. \tag{4}$$

Structural identifiability

Definition of structural identifiability and connection to observability

Partially observed dynamical systems often exhibit structural nonidentifiability. A model is structurally identifiable if a unique parameterization exists for any given model output. A parameter p_i is globally structurally identifiable [3], if for all parameter vectors p, it holds

$$y(p) = y(p') \Rightarrow p_i = p'_i. \tag{5}$$

An individual parameter p_i is structurally nonidentifiable, if changing the parameter does not necessarily alter the model trajectory y, because the changes can be fully compensated by altering other parameters. Local structural identifiability of a parameter is defined by reducing the definition to a neighborhood $v(p)$ instead of the entire parameter space. A model is structurally identifiable, if all of its parameters are structurally identifiable. Multiple related definitions for structural identifiability exist; for a comprehensive discussion, see a recent overview [4].

A structurally nonidentifiable parameter implies the existence of a manifold in parameter space upon which the trajectory y is unchanged. However, on this manifold, the dynamic variables x of the model can change, for example, by a scaling factor, and are thus not uniquely determinable. This is denoted as nonobservability, a concept closely related to parameter nonidentifiability [5–9].

A priori analysis of structural identifiability

Two basic approaches exist to assess structural identifiability of nonlinear dynamic models. *A priori* methods only use the model definition, whereas *a posteriori* methods use the available data to find nonidentifiable parameters. Many *a priori* algorithms have been developed based on a variety of approaches. Powerful methods use Lie group theory because nonidentifiabilities are closely related to symmetries in the system [10–13]. Furthermore, a variety of notable methods exist, which are based on power series expansion [14], generating series [15,16], seminumerical approaches [17,18], differential algebra [19–27], differential geometry [28], and numerical algebraic geometry [29]. For reviews of some of these approaches, see Refs. [28,30,31]. Many of these approaches, especially the early developed methods, can only be applied to rather low-dimensional systems

because of their computational complexity. Thus, recent developments have mainly focused on improving the computational efficiency of the algorithms, for example, by local sensitivity calculations.

As a promising example, Joubert et al. [32] proposed a comprehensive and computationally fast pipeline to cure structural nonidentifiabilities by reparameterization of the model in a five-step procedure: (1) a numerical identifiability analysis based on sensitivities; (2) symbolic identifiability calculations for the low-dimensional candidates from (1), this renders the procedure fast; (3) computation of new model parameters, this step is not unique but requires decisions of the modeler; (4) simplify the original model leading to a lower dimensional parameter vector; and finally (5) check the identifiability of the reparameterized model. In an application to a model with 21 states and 75 parameters, two groups of nonidentifiable parameters were detected, and the model was reparameterized within minutes.

Analysis of structural identifiability using experimental data

In contrast to the aforementioned methods, *a posteriori* methods use the available data to perform identifiability analysis. They infer structural nonidentifiability based on model fits to experimental data. Similar to some sensitivity-based *a priori* approaches, these approaches only assess local structural identifiability.

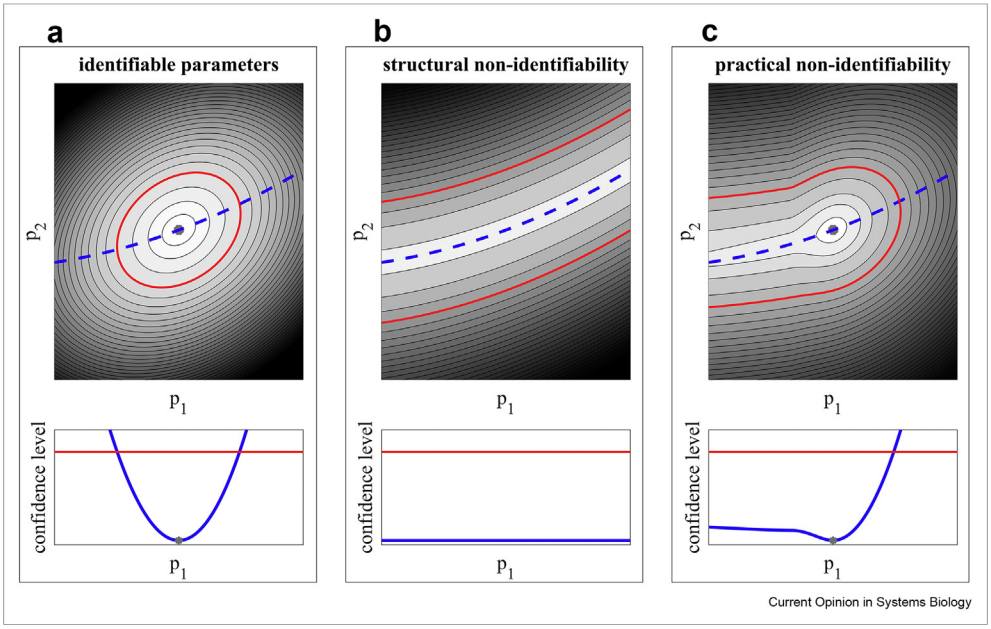
One approach by Hengl et al. [33] suggested to perform numerous fits and investigate nonparametrically whether the final parameter estimates form a low-dimensional manifold in parameter space. This approach also allows to disentangle different sets of coupled nonidentifiable parameters.

An informative and successful method is based on the profile likelihood [34]. The idea of the profile likelihood is to vary one parameter p_i after the other around the maximum likelihood estimate (Equation (4)) and reoptimize the remaining ones.

$$\mathrm{PL}(p_i) = \min_{p_{j\neq i}} \left[\chi^2_{\mathrm{res}}(p)\right] \tag{6}$$

For the two-parameter examples in Figure 1, the blue dashed lines show the path in the parameter space determined by Equation (6). Figure 1A shows the profile likelihood of an identifiable parameter. For a structurally nonidentifiable parameter, the profile likelihood yields a flat line, as shown in Figure 1B. Plotting the remaining

Figure 1



Illustrative example of likelihood contour plots and profile likelihood for an identifiable parameter and structurally and practically nonidentifiable parameters. Subfigures (A), (B), and (C) show contour plots of χ^2_{res} above as well as the profile likelihood versus the parameter below. Lighter colors in the contour plots signify a lower value of χ^2_{res}. Thresholds for confidence intervals corresponding to a confidence level of 95% are shown in red and plotted both in the contour plots and the profile likelihood plots. The lowest value of χ^2_{res} is denoted by a gray asterisk in both the contour plot and the profile likelihood plot. For the identifiable parameter (A), the profile likelihood reaches both an upper and lower threshold, thus leading to a finite confidence interval. For the structurally nonidentifiable parameter (B), the profile likelihood is completely flat, thus yielding infinite confidence intervals. In the contour plot, this translates to a flat path, along which χ^2_{res} does not change. The practically nonidentifiable parameter (C) shows an infinite extension of the low χ^2_{res} region for lower values of the parameter, never reaching the 95% confidence interval threshold. In contrast, a finite upper confidence bound can be derived.

parameters along the profiled parameter reveals which parameters are coupled to the nonidentifiable one [35]. The profile likelihood was recently extended to include two-dimensional profiles to allow for the identification of parameter interdependence [36].

Profile likelihood calculation can be computationally demanding for larger systems because of the numerical reoptimization. Addressing this issue, a fast *a posteriori* method to test identifiability without the need to calculate complete profiles using radial penalization was recently developed [37].

Structural nonidentifiability can also be investigated *a posteriori* by a Bayesian Markov chain Monte Carlo (MCMC) sampling approach. However, for nonidentifiable systems, efficient mixing and thus convergence of the Markov chains is difficult [38]. This problem can be cured by informative priors, but these would mask the problem and should only be implemented if they are based on actual biological insights and prior information. One recent application in the field identified a minimal subset of reactions in a signaling network with a combination of parallel tempering and LASSO regression methods [39].

Reparameterizing structurally nonidentifiable models

Given the recent advances in the computational efficiency of methods, we essentially consider determining structural identifiability no longer a bottleneck in the modeling of nonlinear dynamic systems with ODEs. For models with a high number of connected structurally nonidentifiable parameters, finding and resolving these structural nonidentifiabilities can still be challenging. This is often the case if the number of observed states is much lower than the number of dynamic states. When the structurally nonidentifiable parameters are determined, the problem is usually fixed by a reparameterization of the model. In the simplest case, this is accomplished by fixing some of the involved parameters to a certain value. The price to be paid is typically that the information about the scale of some components is lost, which can limit the predictive power of the model. Nevertheless, biologically meaningful reparameterization of the models after finding nonidentifiabilites remains a challenging task (G. Massonis et al., arXiv:2012.09826v2).

Practical identifiability

From structural to practical identifiability

Structural identifiability implies practical identifiability only for an infinite amount of data with zero noise. Practical identifiability is important for obtaining precise parameter estimates. Moreover, it is especially crucial to ensure that model predictions are well determined. It is analyzed increasingly often to judge a model's predictivity [40–44]. The notion of practical identifiability has been rather vague in the literature, mainly referring to *large confidence intervals* [45–47]. Some approaches exist that define practical identifiability as a combination of model structure and experimental protocol without actual data [48,49]. In contrast, we consider a combination of model and data as practically identifiable if the confidence intervals of all estimated parameters are of finite size [35].

Parameter confidence intervals and identifiability

The profile likelihood (Equation (6)) provides a proper assessment of confidence intervals of estimated parameters in ODE models (Figure 1) by

$$CI_{\mathrm{PL}}(p_i) = \left\{p_i \mid \mathrm{PL}(p_i) \leq \chi^2_{\mathrm{res}}\left(\widehat{p}\right) + \Delta_\alpha\right\}, \qquad (7)$$

where Δ_α denotes the α quantile of the χ^2 distribution with $df = 1$ degrees of freedom for point-wise confidence intervals [34].

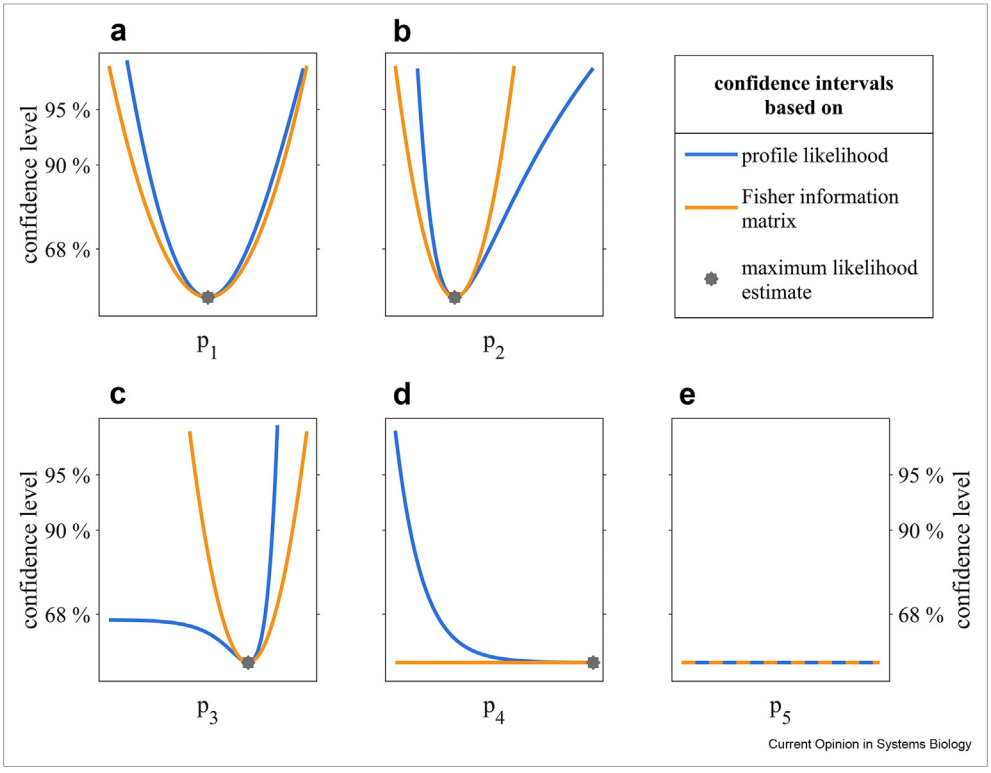
The traditional method for determining confidence intervals based on the Fisher information matrix (FIM) leads to accurate confidence intervals for linear regression models. Because the solutions of all nontrivial ODE models are nonlinear in their parameters, using this method for analyzing, identifiability of such models is questionable [50]. Furthermore, in contrast to FIM-based confidence intervals, profile likelihood–based confidence intervals can be asymmetric and are invariant under reparameterizations of the model, for example, the often applied logarithmic transformation of the parameters. Figure 2 shows five parameters with FIM-based and profile likelihood–based confidence intervals, mainly taken from applications in synthetic biology [51,52].

Identifiability is obtained if all estimated parameters are structurally and practically identifiable, that is, have finite confidence intervals. A nonidentifiable parameter is called practically nonidentifiable if the confidence interval becomes finite for a given confidence level by adding additional measurements for the existing observables (Figure 1C). By adding enough data, a practically nonidentifiable parameter can be made identifiable.

Bayesian methods for identifiability analysis

Bayesian sampling approaches, for example, MCMC, can be used to assess practical identifiability [53–55]. This, however, is only feasible if the model is structurally identifiable because structural nonidentifiabilities will lead to bad mixing of the sampling algorithms. Given a structurally identifiable model, MCMC sampling yields similar results as the profile likelihood analysis [38]. However, a recent application in a spatiotemporal reaction–diffusion model showed that it is one order of magnitude slower than the profile likelihood [56]. To the best of our knowledge, a comprehensive benchmark study comparing the two methods is so far missing.

Figure 2



Parameter confidence intervals based on Fisher information matrix and profile likelihood. Confidence intervals for five parameters based on profile likelihood (blue) and on quadratic approximation using the Fisher information matrix (FIM; orange). FIM-based confidence intervals have two major problems. First, because of the nonlinearity of the underlying systems, the Cramér-Rao bound on the error is invalid, and thus, the FIM-based confidence intervals become uncontrollable for a finite amount of measurements. While in (A), the FIM-based interval is larger than the profile likelihood–based interval, in (B), it is smaller. Second, FIM-based intervals are insensitive to practical nonidentifiabilities. In (C), the FIM-based confidence interval is finite, and thus, the practically nonidentifiable parameter is not detected. In (D), the practically nonidentifiable parameter leads to a flat FIM-based interval, wrongly suggesting structural nonidentifiability. Although the structurally nonidentifiable parameter in (E) is correctly detected, similarly to (D), the calculation of the FIM is challenging because of its singularity in flat likelihood landscapes. The parameters (A)–(D) are adapted from two applications in synthetic biology ((A), (B), (C) from a study by Schneider et al. [52] and (D) from a study by Ochoa-Fernandez et al. [51]). Parameter (E) is from a minimal nonidentifiable toy model. Gray asterisks signify the maximum likelihood estimate of the parameter.

Experimental design and model reduction

Model predictions

To test the predictive power of a model, confidence intervals for the predictions can be computed. For this purpose, forward evaluations of the model are used, for example, bootstrap approaches [57] or sensitivity analysis [58]. They typically require large numerical efforts in the context of nonlinear biological models with a high-dimensional parameter space. A more powerful approach is the *prediction profile likelihood*.

$$\mathrm{PPL}(z) = \min_{p \in \{p \mid g_{\mathrm{pred}}(p) = z\}} \left[\chi^2_{\mathrm{res}}(p)\right], \qquad (8)$$

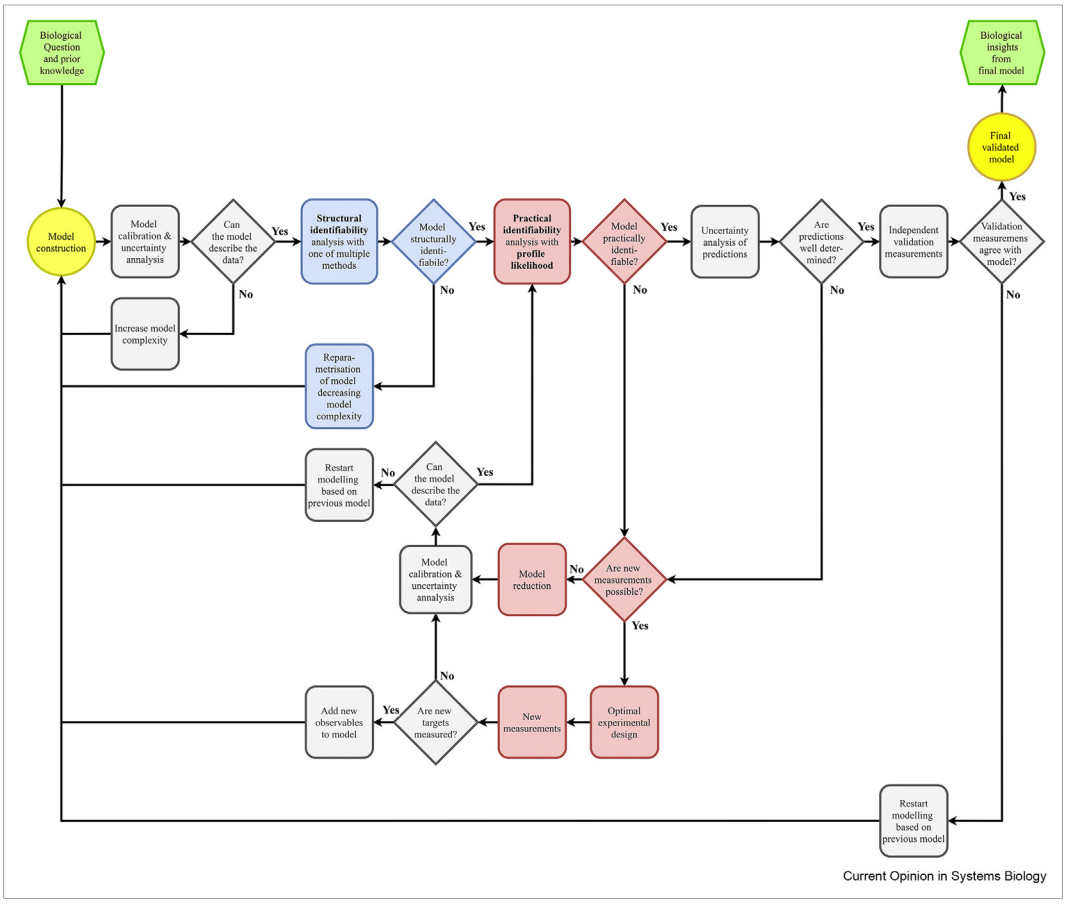
which is obtained by minimizing $\chi^2_{\mathrm{res}}(p)$ (Equation (3)) under the constraint that the model response $g_{\mathrm{pred}}(p)$ is equal to the prediction z. The *prediction profile likelihood* propagates the uncertainty from the experimental data to the prediction by exploring the prediction space instead of the parameter space [59].

Model predictions have to be sufficiently precise to produce insights. For special cases, this can be achieved without identifiability [60,61]. If the model predictions are not of sufficient precision, one has two principal options to tailor the model complexity to the information content of the data: (1) measure additional data, corresponding to an increase of the dimension of the observation function g in Equation (2) or (2) reduce the model complexity according to the available data, corresponding to a decrease of the dimension of the parameter space and/or of the ODE system f in Equation (1). Both options increase the practical identifiability of the model.

Achieving practical identifiability by new measurements with optimal experimental design

Practical identifiability can be achieved by adding new data [44,62]. The process of determining the most informative targets and time points for the new measurements is known as *optimal experimental design* and is frequently applied in different modeling fields, for

Figure 3



Flowchart of the entire modeling process from initial to final model including identifiability analysis. The modeling process begins with the inception of an initial model based on prior knowledge and the underlying biological research question. It ends with the final validated model and the biological insights it provides. The flowchart shows how identifiability analysis is embedded into the overall modeling workflow. The topics discussed in this review related to structural identifiability (blue) and practical identifiability (red) are highlighted with colors in the flowchart. The remaining tiles in gray represent aspects that are beyond the scope of this review. The intricacy of the flowchart shows that the path to biological insights requires multiple iterations of different methods. Identifiability analysis is an integral part of this workflow and should be performed to gain insights from predictive models with well-determined parameters. Furthermore, methods dealing with structural and practical identifiability should always be focused on ultimately progressing along the path toward biological insights.

example, metabolic models [63], animal science [64], linear perturbation networks [65], or synthetic biology [66]. The task is related to the search for an additional measurement that contains the maximal information about the system or parts of it. For improving the identifiability of a specific parameter, the model trajectories along the corresponding parameter profile can be investigated [67,68]. Thereby, measurement points with maximal information content for the parameter of interest can be determined, which corresponds to trajectories with high spread. Similarly, the prediction profile likelihood (Equation (8)) determines the prediction uncertainty of the model at a potential new measurement time point [59], thus promoting the identifiability of the whole model. Measurement points with high prediction uncertainty are effective to constrain the model further, whereas measurements with a low prediction uncertainty are better suited for model selection purposes.

Achieving practical identifiability by reducing model complexity

If measuring additional data is not feasible, the complexity of the model has to be reduced. One way is to fix parameter values or ratios of parameters by means of prior knowledge [69], sensitivity analysis [70,71], or profile likelihood [72]. However, fixing parameters can decrease the interpretative relevance of the model's predictions.

Taking this into account, a systematic model reduction strategy that tailors model complexity to the available data was suggested by Maiwald et al. [73]. On the basis of likelihood profiles, they discuss four basic scenarios

that are discriminated based on the profile likelihood by the combinations of either (1) the profile flattens out for a logarithmized parameter going to infinity or (2) to minus infinity, and either other parameters are (a) coupled to the investigated one or (b) not. For all four possible combinations, there is a cure. For case (1/a), one differential equation is replaced by an algebraic equation; for (1/b), states can be lumped; for (2/a), a variable is fixed, leading to a structural nonidentifiability that can be cured by the methods discussed previously, and for (2/b), a reaction can be removed from the model. This model reduction strategy has been applied, for example, in Refs. [51,52,74]. Independent of the applied method, model reduction steps, and in particular, the conclusions thereof, should always be documented together with the model according to good scientific practice to facilitate reproducibility.

Conclusions

Given the multitude of recently developed methods [13,16,27,32], we consider the file of identifying structurally nonidentifiable parameters as closed. Future research in this field could focus on identifying biologically plausible reparameterizations of the model, for which no comprehensive method yet exists to our knowledge. Furthermore, the extension of the concept of identifiability to different model types, for example, mixed effects models [75,76], is of interest.

Achieving practical identifiability for model and data is more laborious in practice. Practical nonidentifiabilities can be detected reliably, for example, by the profile likelihood method [31]. To achieve identifiability, the model complexity has to be reduced or additional data must be added. Profile likelihood−based model reduction [73] and optimal experimental design [68] provide valuable methods for these purposes. A flowchart locating structural and practical identifiability analysis as discussed in this review within the entire modeling process is given in Figure 3.

Although the availability of advanced methods for the detection and cure of structural and practical nonidentifiabilities is promising, two related challenges remain. In many applications identifiability analysis is not performed with state-of-the-art methods. Particularly, identifiability analysis based on the FIM can be misleading in typical applications in systems biology. We propose a more consequent use of the discussed methods for structural identifiability and especially profile likelihood for practical identifiability analysis to check the limitations and predictive power of mathematical models. In summary, we believe the focal point of research in systems biology should always remain on the biological insights that can be gained from mathematical models, which are structurally and practically identifiable.

Conflict of interest statement

Nothing declared.

Acknowledgements

This work was supported by the German Research Foundation (DFG) under Germany's Excellence Strategy (CIBSS − EXC-2189 − Project ID 390939984), the German Research Foundation (DFG) through grant 272983813/TRR 179, the Deutsche Krebshilfe (grant 70112355), the German Federal Ministry for Education and Research within the research network Systems Medicine of the Liver (LiSyM; grant 031L0048), and by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant INST 35/1134-1 FUGG.
The authors thank Hans Stigter for insightful feedback on the article and Daniel Lill for his help in the writing process.
The graphical abstract and Figure 3 were created with draw.io.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.coisb.2021.03.005.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Box GEP: **Robustness in the strategy of scientific model building**. In *Robustness in statistics*. Edited by Launer RL, Wilkinson GN, New York: Academic Press; 1979:201−236, https://doi.org/10.1016/b978-0-12-438150-6.50018-2.

2. Becker V, Schilling M, Bachmann J, Baumann U, Raue A, Maiwald T, Timmer J, Klingmüller U: **Covering a broad dynamic range: information processing at the erythropoietin receptor**. *Science* 2010, **328**:1404−1408, https://doi.org/10.1126/science.1184913.

3. Walter E, Pronzato L: *Identification of parametric models from experimental data*. London, U.K.: Springer; 1997.

4. * Anstett-Collin F, Denis-Vidal L, Millérioux G: **A priori identifiability: an overview on definitions and approaches**. *Annu Rev Contr* 2020, **50**:139−149, https://doi.org/10.1016/j.arcontrol.2020.10.006.
Rigorously assessing 21 definitions of *a priori* identifiability, the authors provide an overview over analytical and algebraic definitions as well as local and global definitions and their dependencies.

5. Kalman R: **On the general theory of control systems**. *IRE Trans Automatic Control* 1959, **4**:110, https://doi.org/10.1109/TAC.1959.1104873. 110.

6. Bellman R, Åström K: **On structural identifiability**. *Math Biosci* 1970, **7**:329−339, https://doi.org/10.1016/0025-5564(70)90132-X.

7. Chappell MJ, Godfrey KR, Vajda S: **Global identifiability of the parameters of nonlinear systems with specified inputs: a comparison of methods**. *Math Biosci* 1990, **102**:41−73, https://doi.org/10.1016/0025-5564(90)90055-4.

8. * Villaverde AF, Tsiantis N, Banga JR: **Full observability and estimation of unknown inputs, states and parameters of nonlinear biological models**. *J R Soc Interface* 2019, **16**: 20190043, https://doi.org/10.1098/rsif.2019.0043.
A comprehensive approach to observability analysis is provided in this article. This includes assessing inputs, states, and parameters. Structural parameter identifiability is assessed with an approach based on differential geometry.

9. Schmitt T, Ritter B: **Data-based identifiability and observability assessment for nonlinear control systems using the profile likelihood method**. In *21st IFAC world congress, Germany*; 2020, https://doi.org/10.25534/tuprints-00011713.

10. Yates JW, Evans ND, Chappell MJ: **Structural identifiability analysis via symmetries of differential equations**. *Automatica* 2009, **45**:2585−2591, https://doi.org/10.1016/j.automatica.2009.07.009.

11. Merkt B, Kaschek D, Timmer J: **Higher-order Lie symmetries in identifiability and predictability analysis of dynamic models**. *Phys Rev E* 2015, **92**, 012920, https://doi.org/10.1103/PhysRevE.92.012920.

12. Villaverde AF, Evans ND, Chappell MJ, Banga JR: **Input-dependent structural identifiability of nonlinear systems**. *IEEE Contr Syst Lett* 2019, **3**:272–277, https://doi.org/10.1109/LCSYS.2018.2868608.

13. ** Massonis G, Villaverde AF: **Finding and breaking Lie symmetries: implications for structural identifiability and observability in biological modelling**. *Symmetry* 2020, **12**:469, https://doi.org/10.3390/sym12030469.
The authors provide an improved algorithm to find model symmetries using Lie algebra that is computationally more efficient and thus applicable to higher dimensional models.

14. Pohjanpalo H: **System identifiability based on the power series expansion of the solution**. *Math Biosci* 1978, **41**:21–33, https://doi.org/10.1016/0025-5564(78)90063-9.

15. Walter E, Lecourtier Y: **Global approaches to identifiability testing for linear and nonlinear state space models**. *Math Comput Simulat* 1982, **24**:472–482, https://doi.org/10.1016/0378-4754(82)90645-0.

16. * Ligon TS, Fröhlich F, Chiş OT, Banga JR, Balsa-Canto E, Hasenauer J: **GenSSI 2.0: multi-experiment structural identifiability analysis of SBML models**. *Bioinformatics* 2018, **34**:1421–1423, https://doi.org/10.1093/bioinformatics/btx735.
The authors provide an open-source MATLAB toolbox for structural identifiability analysis and the reparametrization of structurally nonidentifiable models. They use the generating series approach and support Systems Biology Markup Language import.

17. Sedoglavic A: **A probabilistic algorithm to test local algebraic observability in polynomial time**. *J Symbolic Comput* 2002, **33**:735–755, https://doi.org/10.1006/jsco.2002.0532.

18. Karlsson J, Anguelova M, Jirstrand M: **An efficient method for structural identifiability analysis of large dynamic systems**. *IFAC Proc* 2012, **45**:941–946, https://doi.org/10.3182/20120711-3-BE-2027.00381.

19. Ljung L, Glad T: **On global identifiability for arbitrary model parametrizations**. *Automatica* 1994, **30**:265–276, https://doi.org/10.1016/0005-1098(94)90029-9.

20. Saccomani MP, Audoly S, D'Angiò L: **Parameter identifiability of nonlinear systems: the role of initial conditions**. *Automatica* 2003, **39**:619–632, https://doi.org/10.1016/S0005-1098(02)00302-3.

21. Bellu G, Saccomani MP, Audoly S, D'Angiò L: **DAISY: a new software tool to test global identifiability of biological and physiological systems**. *Comput Methods Progr Biomed* 2007, **88**:52–61, https://doi.org/10.1016/j.cmpb.2007.07.002.

22. Meshkat N, Kuo CE-Z, DiStefano J: **On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and COMBOS: a novel web implementation**. *PloS One* 2014, **9**, e110261, https://doi.org/10.1371/journal.pone.0110261.

23. Thomaseth K, Saccomani MP: **Local identifiability analysis of nonlinear ODE models: how to determine all candidate solutions**. *IFAC-PapersOnLine* 2018, **51**:529–534, https://doi.org/10.1016/j.ifacol.2018.03.089.

24. Varghese A, Narasimhan S, Bhatt N: **A priori parameter identifiability in complex reaction networks**. *IFAC-PapersOnLine* 2018, **51**:760–765, https://doi.org/10.1016/j.ifacol.2018.09.162.

25. Saccomani MP, Bellu G, Audoly S, D'Angió L: **A new version of DAISY to test structural identifiability of biological models**. In *Computational methods in systems biology*; 2019:329–334, https://doi.org/10.1007/978-3-030-31304-3_21.

26. Hong H, Ovchinnikov A, Pogudin G, Yap C: **SIAN: software for structural identifiability analysis of ODE models**. *Bioinformatics* 2019, **35**:2873–2874, https://doi.org/10.1093/bioinformatics/bty1069.

27. ** Hong H, Ovchinnikov A, Pogudin G, Yap C: **Global identifiability of differential models**. *Commun Pure Appl Math* 2020, **73**:1831–1879, https://doi.org/10.1002/cpa.21921.
Combining a probabilistic and analytical criterion in one algorithm, the authors perform *a priori* structural identifiability analysis efficiently and quickly. The algorithm is available in the SIAN Maple software package.

28. Villaverde AF: **Observability and structural identifiability of nonlinear biological systems**. *Complexity* 2019, **2019**:1–12, https://doi.org/10.1155/2019/8497093.

29. Bates DJ, Hauenstein JD, Meshkat N: **Identifiability and numerical algebraic geometry**. *PloS One* 2019, **14**, e0226299, https://doi.org/10.1371/journal.pone.0226299.

30. Chis OT, Banga JR, Balsa-Canto E: **Structural identifiability of systems biology models: a critical comparison of methods**. *PloS One* 2011, **6**, e27755, https://doi.org/10.1371/journal.pone.0027755.

31. Raue A, Karlsson J, Saccomani MP, Jirstrand M, Timmer J: **Comparison of approaches for parameter identifiability analysis of biological systems**. *Bioinformatics* 2014, **30**:1440–1448, https://doi.org/10.1093/bioinformatics/btu006.

32. ** Joubert D, Stigter JD, Molenaar J: **An efficient procedure to assist in the re-parametrization of structurally unidentifiable models**. *Math Biosci* 2020, **323**:108328, https://doi.org/10.1016/j.mbs.2020.108328.
Using a combination of numerical and algebraic methods, the authors provide an incredibly fast algorithm for assessing structural identifiability. Their approach uses sensitivity analysis to provide candidates for further symbolic analysis, before computing new algebraic relations for the nonidentifiable parameters.

33. Hengl S, Kreutz C, Timmer J, Maiwald T: **Data-based identifiability analysis of non-linear dynamical models**. *Bioinformatics* 2007, **23**:2612–2618, https://doi.org/10.1093/bioinformatics/btm382.

34. Murphy SA, van der Vaart AW: **On profile likelihood**. *J Am Stat Assoc* 2000, **95**:449, https://doi.org/10.2307/2669386.

35. Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J: **Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood**. *Bioinformatics* 2009, **25**:1923–1929, https://doi.org/10.1093/bioinformatics/btp358.

36. * Brastein OM, Lie B, Sharma R, Skeie N-O: **Parameter estimation for externally simulated thermal network models**. *Energy Build* 2019, **191**:200–210, https://doi.org/10.1016/j.enbuild.2019.03.018.
Extension of the profile likelihood method for 2D likelihood profiles that allow for analysis of pairwise parameter interdependence.

37. ** Kreutz C: **An easy and efficient approach for testing identifiability**. *Bioinformatics* 2018, **34**:1913–1921, https://doi.org/10.1093/bioinformatics/bty035.
Fast alternative method to test structural and practical identifiability without the need to calculate the profile likelihood for each model parameter.

38. Raue A, Kreutz C, Theis FJ, Timmer J: **Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability**. *Phil Trans R Soc A* 2013, **371**:20110544, https://doi.org/10.1098/rsta.2011.0544.

39. Gupta S, Lee REC, Faeder JR: **Parallel tempering with LASSO for model reduction in systems biology**. *PLoS Comput Biol* 2020, **16**, e1007669, https://doi.org/10.1371/journal.pcbi.1007669.

40. Busch K, Klapproth K, Barile M, Flossdorf M, Holland-Letz T, Schlenner SM, Reth M, Höfer T, Rodewald H-R: **Fundamental properties of unperturbed haematopoiesis from stem cells in vivo**. *Nature* 2015, **518**:542–546, https://doi.org/10.1038/nature14242.

41. David I, Ricard A: **A unified model for inclusive inheritance in livestock species**. *Genetics* 2019, **212**:1075–1099, https://doi.org/10.1534/genetics.119.302375.

42. Duchesne R, Guillemin A, Crauste F, Gandrillon O: **Calibration, selection and identifiability analysis of a mathematical model of the in vitro erythropoiesis in normal and perturbed contexts**. *Silico Biol* 2019, **13**:55–69, https://doi.org/10.3233/ISB-190471.

43. Zhou W, Huang R, Liu K, Zhang W: **A novel interval-based approach for quantifying practical parameter identifiability of**

a lithium-ion battery model. *Int J Energy Res* 2020, **44**: 3558–3573, https://doi.org/10.1002/er.5118.

44. * Johnson KE, Howard GR, Morgan D, Brenner EA, Gardner AL, Durrett RE, Mo W, Al'Khafaji A, Sontag ED, Jarrett AM, Yankeelov TE, Brock A: **Integrating transcriptomics and bulk time course data into a mathematical framework to describe and predict therapeutic resistance in cancer**. *Phys Biol* 2020, **18**, 016001, https://doi.org/10.1088/1478-3975/abb09c.
The authors emphasize practical identifiability as a central requirement for the validity of their model predictions. It is achieved by inclusion of single-cell RNA-Seq data into the ODE model via a machine learning algorithm.

45. Nihtilä M, Virkkunen J: **Practical identifiability of growth and substrate consumption models**. *Biotechnol Bioeng* 1977, **19**: 1831–1850, https://doi.org/10.1002/bit.260191208.

46. Holmberg A: **On the practical identifiability of microbial growth models incorporating Michaelis-Menten type nonlinearities**. *Math Biosci* 1982, **62**:23–43, https://doi.org/10.1016/0025-5564(82)90061-X.

47. Miao H, Xia X, Perelson AS, Wu H: **On identifiability of nonlinear ODE models and applications in viral dynamics**. *SIAM Rev* 2011, **53**:3–39, https://doi.org/10.1137/090757009.

48. * Gontier C, Pfister J-P: **Identifiability of a binomial synapse**. *Front Comput Neurosci* 2020, **14**, https://doi.org/10.3389/fncom.2020.558477.
An alternative definition of practical identifiability is provided by Gontier and Pfister, which is independent of the specific data realization. This method allows to study how practical identifiability is affected by different experimental protocols. However, it requires a nested hierarchy of competing models, which is not clear *a priori* for complex models as used frequently in systems biology.

49. Saccomani MP, Thomaseth K: **The union between structural and practical identifiability makes strength in reducing oncological model complexity: a case study**. *Complexity* 2018, **2018**:1–10, https://doi.org/10.1155/2018/2380650.

50. Neale MC, Miller MB: **The use of likelihood-based confidence intervals in genetic models**. *Behav Genet* 1997, **27**:113–120, https://doi.org/10.1023/A:1025681223921.

51. Ochoa-Fernandez R, Abel NB, Wieland F-G, Schlegel J, Koch L-A, Miller JB, Engesser R, Giuriani G, Brandl SM, Timmer J, Weber W, Ott T, Simon R, Zurbriggen MD: **Optogenetic control of gene expression in plants in the presence of ambient white light**. *Nat Methods* 2020, **17**:717–725, https://doi.org/10.1038/s41592-020-0868-y.

52. Schneider N, Wieland F-G, Kong D, Fischer AAM, Hörner M, Timmer J, Ye H, Weber W: **Liquid-liquid phase separation of light-inducible transcription factors increases transcription activation in mammalian cells and mice**. *Sci Adv* 2021, **7**, eabd3568, https://doi.org/10.1126/sciadv.abd3568.

53. Siekmann I, Sneyd J, Crampin EJ: **MCMC can detect nonidentifiable models**. *Biophys J* 2012, **103**:2275–2286, https://doi.org/10.1016/j.bpj.2012.10.024.

54. Hines KE, Middendorf TR, Aldrich RW: **Determination of parameter identifiability in nonlinear biophysical models: a Bayesian approach**. *J Gen Physiol* 2014, **143**:401–416, https://doi.org/10.1085/jgp.201311116.

55. Zuo Y, Sarkar S, Corwin MT, Olson K, Badawi RD, Wang G: **Structural and practical identifiability of dual-input kinetic modeling in dynamic PET of liver inflammation**. *Phys Med Biol* 2019, **64**:175023, https://doi.org/10.1088/1361-6560/ab1f29.

56. * Simpson MJ, Baker RE, Vittadello ST, Maclaren OJ: **Practical parameter identifiability for spatio-temporal models of cell invasion**. *J R Soc Interface* 2020, **17**:20200055, https://doi.org/10.1098/rsif.2020.0055.
By means of spatiotemporal models, the authors show that profile likelihood leads to the same results as the established MCMC approach while being one order of magnitude faster to compute.

57. Joshi M, Seidel-Morgenstern A, Kremling A: **Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems**. *Metab Eng* 2006, **8**:447–455, https://doi.org/10.1016/j.ymben.2006.04.003.

58. Sachs L: *Applied statistics, springer series in statistics*. New York, New York, NY: Springer; 1984, https://doi.org/10.1007/978-1-4612-5246-7.

59. Kreutz C, Raue A, Timmer J: **Likelihood based observability analysis and confidence intervals for predictions of dynamic models**. *BMC Syst Biol* 2012, **6**:120, https://doi.org/10.1186/1752-0509-6-120.

60. Cedersund G: **Prediction uncertainty estimation despite unidentifiability: an overview of recent developments**. In *Uncertainty in biology*; 2016:449–466, https://doi.org/10.1007/978-3-319-21296-8_17.

61. Rateitschak K, Winter F, Lange F, Jaster R, Wolkenhauer O: **Parameter identifiability and sensitivity analysis predict targets for enhancement of STAT1 activity in pancreatic cancer and stellate cells**. *PLoS Comput Biol* 2012, **8**, e1002815, https://doi.org/10.1371/journal.pcbi.1002815.

62. Bhonsale SS, Stokbroekx B, Van Impe J: **Assessment of the parameter identifiability of population balance models for air jet mills**. *Comput Chem Eng* 2020, **143**:107056, https://doi.org/10.1016/j.compchemeng.2020.107056.

63. Frøysa HG, Skaug HJ, Alendal G: **Experimental design for parameter estimation in steady-state linear models of metabolic networks**. *Math Biosci* 2020, **319**:108291, https://doi.org/10.1016/j.mbs.2019.108291.

64. Muñoz-Tamayo R, Puillet L, Daniel JB, Sauvant D, Martin O, Taghipoor M, Blavy P: **Review: to be or not to be an identifiable model. Is this a relevant question in animal science modelling?** *Animal* 2018, **12**:701–712, https://doi.org/10.1017/S1751731117002774.

65. Gross T, Blüthgen N: **Identifiability and experimental design in perturbation studies**. *Bioinformatics* 2020, **36**:i482–i489, https://doi.org/10.1093/bioinformatics/btaa404.

66. Bandiera L, Gomez-Cabeza D, Gilman J, Balsa-Canto E, Menolascina F: **Optimally designed model selection for synthetic biology**. *ACS Synth Biol* 2020, **9**:3134–3144, https://doi.org/10.1021/acssynbio.0c00393.

67. Raue A, Becker V, Klingmüller U, Timmer J: **Identifiability and observability analysis for experimental design in nonlinear dynamical models**. *Chaos* 2010, **20**:45105, https://doi.org/10.1063/1.3528102.

68. Steiert B, Raue A, Timmer J, Kreutz C: **Experimental design for parameter estimation of gene regulatory networks**. *PloS One* 2012, **7**, e40052, https://doi.org/10.1371/journal.pone.0040052.

69. Tivay A, Jin X, Lo AK-Y, Scully CG, Hahn J-O: **Practical use of regularization in individualizing a mathematical model of cardiovascular hemodynamics using scarce data**. *Front Physiol* 2020, **11**, https://doi.org/10.3389/fphys.2020.00452.

70. He QP, Wang J: **Application of systems engineering principles and techniques in biological big data analytics: a review**. *Processes* 2020, **8**:951, https://doi.org/10.3390/pr8080951.

71. Zi Z: **Sensitivity analysis approaches applied to systems biology models**. *IET Syst Biol* 2011, **5**:336–346, https://doi.org/10.1049/iet-syb.2011.0015.

72. Pironet A, Docherty PD, Dauby PC, Chase JG, Desaive T: **Practical identifiability analysis of a minimal cardiovascular system model**. *Comput Methods Progr Biomed* 2019, **171**: 53–65, https://doi.org/10.1016/j.cmpb.2017.01.005.

73. Maiwald T, Hass H, Steiert B, Vanlier J, Engesser R, Raue A, Kipkeew F, Bock HH, Kaschek D, Kreutz C, Timmer J: **Driving the model to its limit: profile likelihood based model reduction**. *PloS One* 2016, **11**, e0162366, https://doi.org/10.1371/journal.pone.0162366.

74. * Tönsing C, Timmer J, Kreutz C: **Profile likelihood-based analyses of infectious disease models**. *Stat Methods Med Res* 2018, **27**:1979–1998, https://doi.org/10.1177/0962280217746444.
This study applies profile likelihood–based identifiability analysis to infectious diseases models, which are based on ODE models similarly to systems biology context. Using profile likelihood, a stepwise data-based model reduction procedure is shown for a Zika virus model.

75. Cucurull-Sanchez L, Chappell MJ, Chelliah V, Amy Cheung SY, Derks G, Penney M, Phipps A, Malik-Sheriff RS, Timmis J, Tindall MJ, Graaf PH, Vicini P, Yates JWT: **Best practices to maximize the use and reuse of quantitative and systems pharmacology models: recommendations from the United Kingdom quantitative and systems pharmacology network**. *CPT Pharmacometrics Syst Pharmacol* 2019, **8**:259–272, https://doi.org/10.1002/psp4.12381.

76. Janzén DL, Jirstrand M, Chappell MJ, Evans ND: **Three novel approaches to structural identifiability analysis in mixed-effects models**. *Comput Methods Progr Biomed* 2019, **171**:141–152, https://doi.org/10.1016/j.cmpb.2016.04.024.