On structural and practical identifiability

Franz-Georg Wieland[a,b,c], Adrian L. Hauber[a,b], Marcus Rosenblatt[a,b], Christian Tönsing[a,b,c], Jens Timmer[a,b,c,*]

[a]*Institute of Physics, University of Freiburg, Hermann-Herder-Str. 3, 79104 Freiburg, Germany*
[b]*Freiburg Center for Data Analysis and Modelling (FDM), University of Freiburg, Ernst-Zermelo-Str. 1, 79104 Freiburg, Germany*
[c]*Centre for Integrative Biological Signalling Studies (CIBSS), University of Freiburg, Schänzlestr. 18, 79104 Freiburg, Germany*

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 non-identifiability has a long tradition. Computationally efficient methods to detect and cure it have been developed recently. Practical non-identifiability 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 modelling 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 non-identifiability.

Highlights

- With recent advances structural identifiability is no longer a major issue
- Practical identifiability is still challenging
- Fisher information matrix is misleading
- Profile likelihood can solve practical identifiability challenge

Keywords: Identifiability, Structural identifiability, Practical identifiability, Profile likelihood, Fisher information matrix, Non-linear dynamics, ODE models, Experimental design, Model reduction, Observability

arXiv:2102.05100v1 [stat.ME] 9 Feb 2021

Graphical abstract
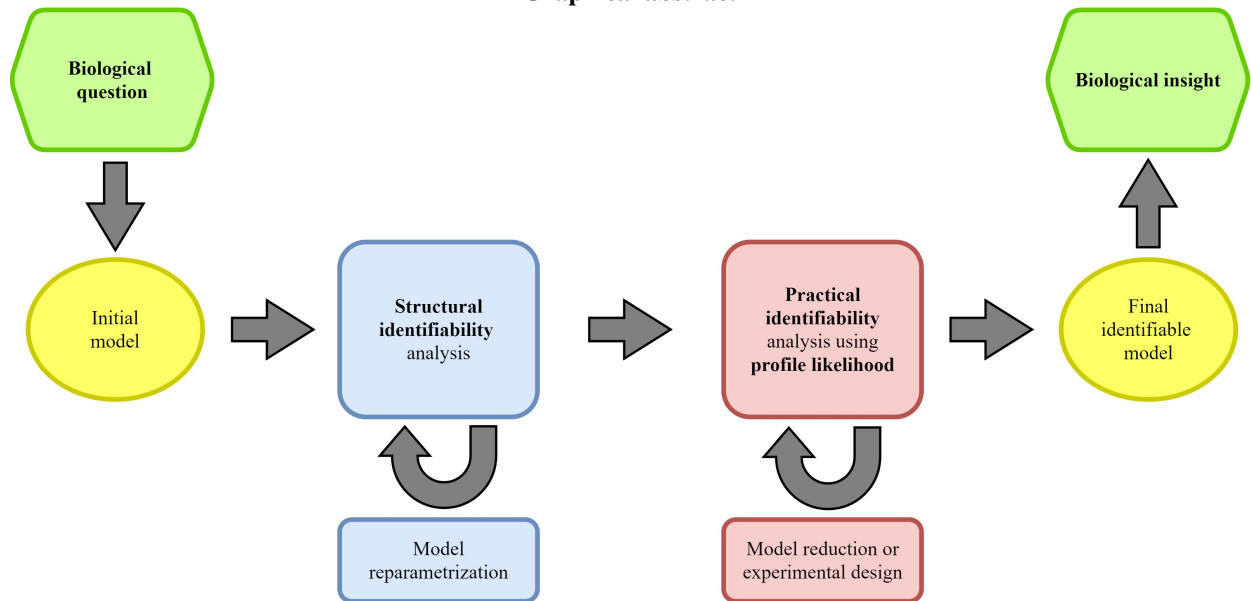
Introduction

Biological modelling and Box's statement

Traditional biological reasoning often is rather qualitative, descriptive, and static, which results e.g. 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 modelling in general, in 1979, George E.P. Box coined his famous statement: "All models are wrong, but some are useful" [1]. While the former part of the quote is intuitively clear, since 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. Firstly, it has to capture the main effects of the question of interest, i.e. describe the data with reasonable accuracy, and neglect the rest. Secondly, a *useful* model has to make experimentally falsifiable predictions in order to be testable. Models that exhibit these two properties are *good* models. Thirdly, the model should enable to gather insights about the biological system. In a typical modelling 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. Based on biological intuition and trial-and-error, one increases the model complexity until the data can

be fitted. Often, this leads to an over-parameterised model that over-fits 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 towards a *good* model is laborious: additional data needs 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 modelling 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 modelling biological systems because the limited amount and quality of the experimental data with large measurement noise in only partially observed systems often leads to *bad* models during the modelling process. Concerning identifiability, one distinguishes between structural identifiability dealing with inherently indeterminable parameters due to 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 $u(t)$. Since 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, e.g. because of the limited availability or restricted capability of antibodies to discriminate between un-phosphorylated, i.e. inactive, and phosphorylated, i.e. active, proteins. Thus, an observation function $g(\cdot)$ is required that 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, non-linear, sparse and non-autonomous, 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

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

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

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

Structural identifiability

Definition of structural identifiability and connection to observability

Partially observed dynamical systems often exhibit structural non-identifiability. Structural identifiability is the ability to uniquely estimate parameters from any given model output. A parameter p_i is globally structurally identifiable [3], if for all parameter vectors p, it holds

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

An individual parameter p_i is structurally non-identifiable, if changing the parameter does not 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 neighbourhood $v(p)$ 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 non-identifiable 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, e.g. by a scaling factor, and are thus not uniquely determinable. This is denoted as non-observability, a concept closely related to parameter non-identifiability [5, 6, 7, 8, 9].

A priori analysis of structural identifiability

Two basic approaches exist to assess structural identifiability of non-linear dynamic models. *A priori* methods only use the model definition, while *a posteriori* methods use the available data to find non-identifiable parameters. Many *a priori* algorithms have been developed based on a variety of approaches. Powerful methods use Lie group theory, since non-identifiabilities are closely related to symmetries in the system [10, 11, 12, 13]. Furthermore, a variety of notable methods exist, which are based on power series expansion [14], generating series [15, 16], semi-numerical approaches [17, 18], differential algebra [19, 20, 21, 22, 23, 24, 25, 26, 27], differential geometry [28],

and numerical algebraic geometry [29]. For reviews of some of these approaches, see [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, e.g. by local sensitivity calculations.

As a promising example, Joubert et al. [32] proposed a comprehensive and computationally fast pipeline to cure structural non-identifiabilities by re-parameterisation of the model in a five-step procedure: (i) a numerical identifiability analysis based on sensitivities, (ii) symbolic identifiability calculations for the low-dimensional candidates from (i), this renders the procedure fast, (iii) computation of new model parameters, this step is not unique, but requires decisions of the modeler, (iv) simplify the original model leading to a lower dimensional parameter vector, and finally (v) check the identifiability of the re-parameterised model. In an application to a model with 21 states and 75 parameters, two groups of non-identifiable parameters were detected and the model was re-parameterised 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 non-identifiability 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 non-parametrically whether the final parameter estimates form a low-dimensional manifold in parameter space. This approach also allows to disentangle different sets of coupled non-identifiable 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 re-optimise the remaining ones

$$\mathrm{PL}\left(p_i\right) = \min_{p_{j \neq i}} \left[\chi^2_{\mathrm{res}}\left(p\right)\right].$$ (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 non-identifiable parameter the profile likelihood yields a flat line as shown in Figure 1B. Plotting the remaining parameters along the profiled parameter reveals which parameters are coupled to the non-identifiable 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 due to the numerical re-optimisation. Addressing this issue, a fast *a posteriori* method to test identifiability without the need to calculate complete profiles using radial penalisation was recently developed [37].
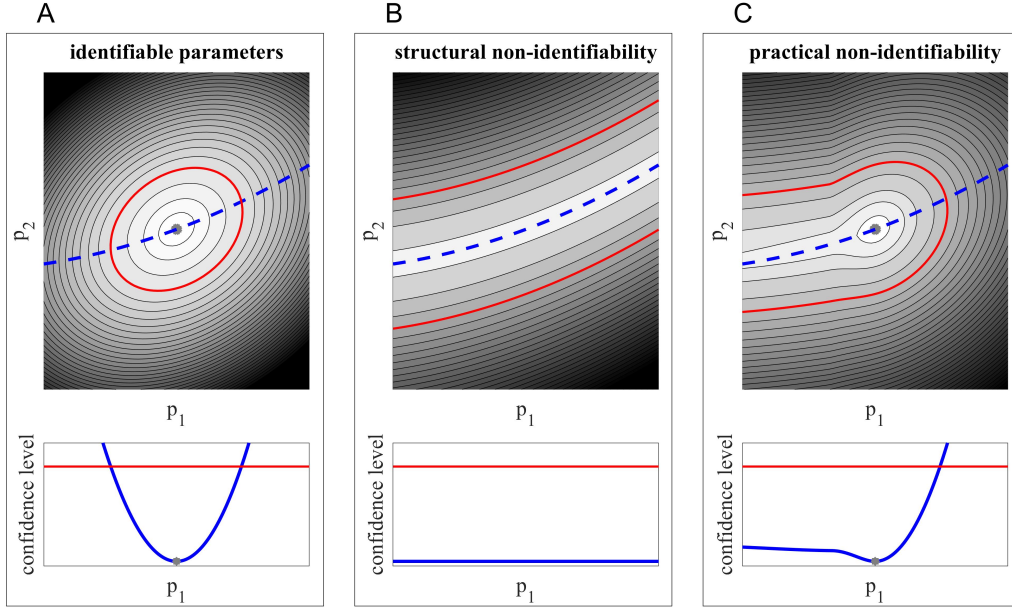
Figure 1: **Likelihood contour plots and profile likelihood for an identifiable parameter and structurally and practically non-identifiable parameters.** Subfigures (A), (B), and (C) show contour plots of χ^2_{res} above as well as the profile likelihood versus the parameter below. Lighter colours in 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 grey 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 non-identifiable 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 non-identifiable 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.

Structural non-identifiability can also be investigated *a posteriori* by a Bayesian Markov chain Monte Carlo (MCMC) sampling approach. However, for non-identifiable 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 signalling network with a combination of parallel tempering and LASSO regression methods [39].

Re-parameterising structurally non-identifiable models

Given the recent advances in the computational efficiency of methods, we essentially consider determining structural identifiability no longer a bottleneck in the modelling of non-linear dynamic systems with ODEs. When the structurally non-identifiable parameters are determined, the problem is usually fixed by a re-parameterisation 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. Nevertheless, biolog-

ically meaningful re-parameterisation of the models after finding non-identifiabilites 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, 41, 42, 43, 44]. The notion of practical identifiability has been rather vague in the literature, mainly referring to *large confidence intervals* [45, 46, 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

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

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. Since the solutions of all nontrivial ODE models are non-linear in their parameters, using this method for analysing 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 re-parameterisations of the model, e.g. the often applied logarithmic transformation of the parameters. Figure 2 shows five parameters with FIM- 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, i.e. have finite confidence intervals. A non-identifiable parameter is called practically non-identifiable if its confidence interval can be narrowed by adding additional measurements for the existing observables and experimental conditions (Figure 1C).
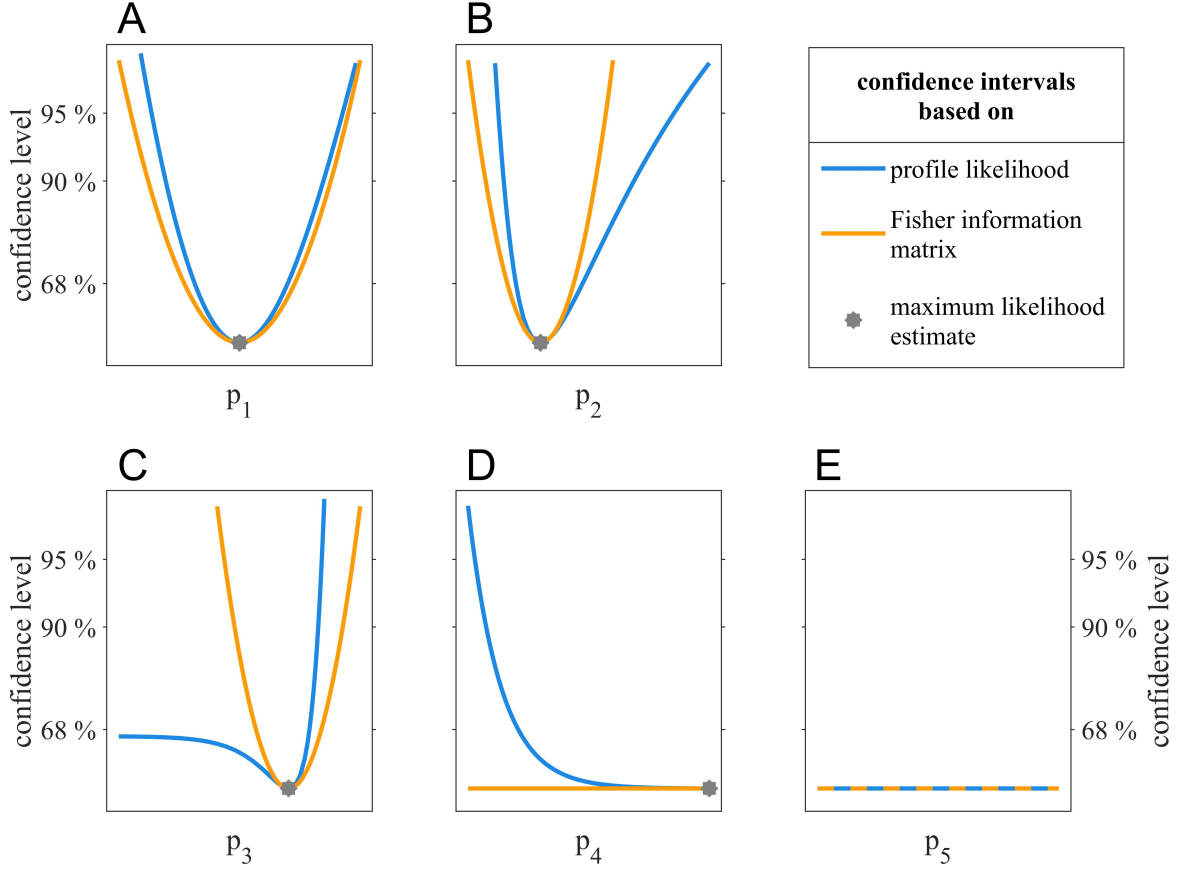
7

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. Firstly, due to the non-linearity 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. Secondly, FIM based intervals are insensitive to practical non-identifiabilities. In (C), the FIM-based confidence interval is finite and thus the practically non-identifiable parameter is not detected. In (D), the practically non-identifiable parameter leads to a flat FIM-based interval, wrongly suggesting structural non-identifiability. While the structurally non-identifiable parameter in (E) is correctly detected, similarly to (D) the calculation of the FIM is challenging due to its singularity in flat likelihood landscapes. The parameters (A)-(D) are adapted from two applications in synthetic biology ((A),(B),(C) from Schneider et al. [52], (D) from Ochoa-Fernandez et al. [51]). Parameter (E) is from a minimal non-identifiable toy model. Grey asterisks signify the maximum likelihood estimate of the parameter.

Bayesian methods for identifiability analysis

Bayesian sampling approaches, e.g. MCMC, can be used to assess practical identifiability [53, 54, 55]. This, however is only feasible if the model is structurally identifiable, since structural non-identifiabilities 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 spatio-temporal reaction–diffusion model

showed, that it is one order of magnitude slower than the profile likelihood [56].

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 utilized, e.g. bootstrap approaches [57] or sensitivity analysis [58]. They typically require large numerical efforts in the context of non-linear biological models with a high-dimensional parameter space. A more powerful approach is the *prediction profile likelihood*

$$\text{PPL}(z) = \min_{p \in \{p \mid g_{\text{pred}}(p) = z\}} \left[\chi_{\text{res}}^2 (p) \right],\tag{8}$$

which is obtained by minimising $\chi_{\text{res}}^2(p)$ (Equation (3)) under the constraint that the model response $g_{\text{pred}}(p)$ 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].

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: (i) measure additional data, corresponding to an increase of the dimension of the observation function g in Equation (2), or (ii) 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, 60]. 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 modelling fields, e.g. metabolic models [61], animal science [62], linear perturbation networks [63] or synthetic biology [64]. The task is related to the search of 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 [65, 66]. 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 [67], sensitivity analysis [68, 69] or profile likelihood [70]. 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. [71]. Based on likelihood profiles, they discuss four basic scenarios that are discriminated based on the profile likelihood by the combinations of: either (i) the profile flattens out for a logarithmised parameter going to infinity or (ii) 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 (i/a), one differential equation is replaced by an algebraic equation, for (i/b), states can be lumped, for (ii/a), a variable is fixed, leading to a structural non-identifiability that can be cured by the methods discussed above, and for (ii/b), a reaction can be removed from the model. This model reduction strategy has been applied e.g. in [51, 52, 72]. 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 non-identifiable parameters as closed. Future research in this field could focus on identifying biologically plausible re-parameterisations of the model, for which no comprehensive method yet exists. Furthermore, the extension of the concept of identifiability to different model types, e.g. mixed effects models [73, 74], is of interest.

Achieving practical identifiability for model and data is more laborious in practice. Practical non-identifiabilities can be detected reliably, e.g. by the profile likelihood method [31]. In order to achieve identifiability, the model complexity has to be reduced or additional data must be added. Profile likelihood-based model reduction [71] and optimal experimental design [66] provide valuable methods for these purposes. A flowchart locating structural and practical identifiability analysis as discussed in this review within the entire modelling process is given in Figure 3.

Although the availability of advanced methods for the detection and cure of structural and practical non-identifiabilities 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 Fisher information matrix 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 in order 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.
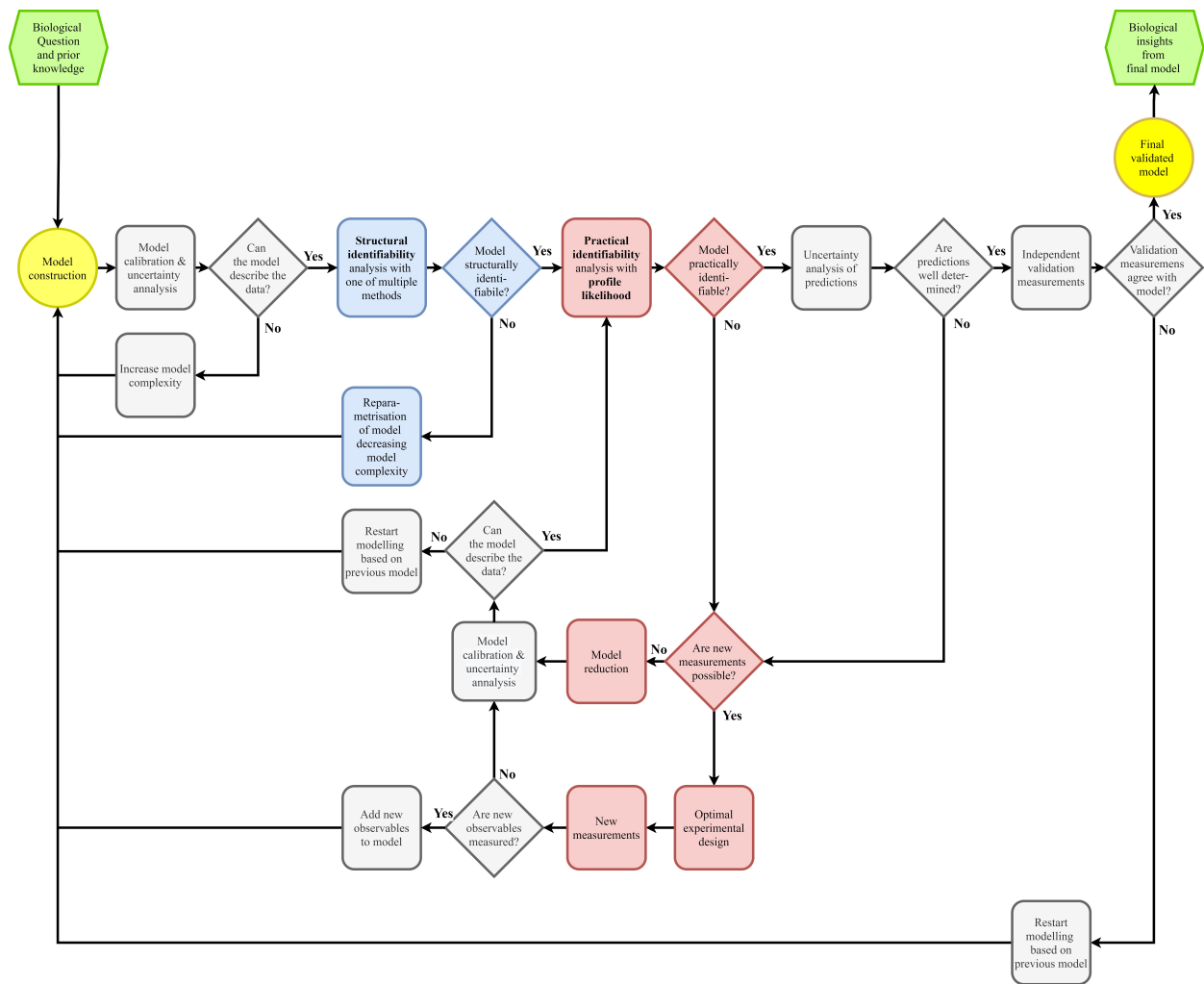
Figure 3: **Flowchart of the entire modelling process from inital to final model including identifiability analysis.** The modelling process begins with the inception of an initial model based on prior knowledge and the underlying biological research question. It ends with the final validated model and the biological insights it provides. The flowchart shows, how identifiability analysis is embedded into the overall modelling workflow. The topics discussed in this review related to structural identifiability (blue) and practical identifiability (red) are highlighted with colours 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 towards biological insights.

Conflict of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

[1] G. E. P. Box, Robustness in the strategy of scientific model building, in: R. L. Launer, G. N. Wilkinson (Eds.), Robustness in Statistics, Academic Press, New York, 1979, pp. 201–236. doi:10.1016/b978-0-12-438150-6.50018-2.

[2] V. Becker, M. Schilling, J. Bachmann, U. Baumann, A. Raue, T. Maiwald, J. Timmer, U. Klingmüller, Covering a broad dynamic range: Information processing at the erythropoietin receptor, Science 328 (2010) 1404–1408. doi:10.1126/science.1184913.

[3] E. Walter, L. Pronzato, Identification of Parametric Models from Experimental Data, Springer, London, U.K., 1997.

[4] F. Anstett-Collin, L. Denis-Vidal, G. Millérioux, A priori identifiability: An overview on definitions and approaches, Annual Reviews in Control 50 (2020) 139–149. doi: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] R. Kalman, On the general theory of control systems, IRE Transactions on Automatic Control 4 (1959) 110–110. doi:10.1109/TAC.1959.1104873.

[6] R. Bellman, K. Åström, On structural identifiability, Mathematical Biosciences 7 (1970) 329–339. doi:10.1016/0025-5564(70)90132-X.

[7] M. J. Chappell, K. R. Godfrey, S. Vajda, Global identifiability of the parameters of nonlinear systems with spezified inputs: a comparison of methods, Math. Biosciences 102 (1990) 41–73. doi:10.1016/0025-5564(90)90055-4.

[8] A. F. Villaverde, N. Tsiantis, J. R. Banga, Full observability and estimation of unknown inputs, states and parameters of nonlinear biological models, Journal of The Royal Society Interface 16 (2019) 20190043. doi:10.1098/rsif.2019.0043.

   *A comprehensive approach to observability analysis is provided in this paper. This includes assessing inputs, states and parameters. Structural parameter identifiability is assessed with an approach based on differential geometry.

[9] T. Schmitt, B. Ritter, Data-based identifiability and observability assessment for nonlinear control systems using the profile likelihood method, in: 21st IFAC World Congress, Germany, 2020. doi:10.25534/tuprints-00011713.

[10] J. W. Yates, N. D. Evans, M. J. Chappell, Structural identifiability analysis via symmetries of differential equations, Automatica 45 (2009) 2585–2591. doi:10.1016/j.automatica.2009.07.009.

[11] B. Merkt, D. Kaschek, J. Timmer, Higher-order Lie symmetries in identifiability and predictability analysis of dynamic models, Physical Review E 92 (2015) 012920. doi:10.1103/PhysRevE.92.012920.

[12] A. F. Villaverde, N. D. Evans, M. J. Chappell, J. R. Banga, Input-dependent structural identifiability of nonlinear systems, IEEE Control Systems Letters 3 (2019) 272–277. doi:10.1109/LCSYS.2018.2868608.

[13] G. Massonis, A. F. Villaverde, Finding and breaking Lie symmetries: Implications for structural identifiability and observability in biological modelling, Symmetry 12 (2020) 469. doi: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] H. Pohjanpalo, System identifiability based on the power series expansion of the solution, Mathematical Biosciences 41 (1978) 21–33. doi:`10.1016/0025-5564(78)90063-9`.

[15] E. Walter, Y. Lecourtier, Global approaches to identifiability testing for linear and nonlinear state space models, Mathematics and Computers in Simulation 24 (1982) 472–482. doi:`10.1016/0378-4754(82)90645-0`.

[16] T. S. Ligon, F. Fröhlich, O. T. Chiş, J. R. Banga, E. Balsa-Canto, J. Hasenauer, GenSSI 2.0: Multi-experiment structural identifiability analysis of SBML models, Bioinformatics 34 (2018) 1421–1423. doi:`10.1093/bioinformatics/btx735`.

    *The authors provide an open-source Matlab toolbox for structural identifiability analysis and the reparametrization of structurally non-identifiable models. They use the generating series approach and support Systems Biology Markup Language import.

[17] A. Sedoglavic, A probabilistic algorithm to test local algebraic observability in polynomial time, Journal of Symbolic Computation 33 (2002) 735–755. doi:`10.1006/jsco.2002.0532`.

[18] J. Karlsson, M. Anguelova, M. Jirstrand, An efficient method for structural identifiability analysis of large dynamic systems, IFAC Proceedings Volumes 45 (2012) 941–946. doi:`10.3182/20120711-3-BE-2027.00381`.

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

[20] M. P. Saccomani, S. Audoly, L. D'Angiò, Parameter identifiability of nonlinear systems: The role of initial conditions, Automatica 39 (2003) 619–632. doi:`10.1016/S0005-1098(02)00302-3`.

[21] G. Bellu, M. P. Saccomani, S. Audoly, L. D'Angiò, DAISY: A new software tool to test global identifiability of biological and physiological systems, Computer Methods and Programs in Biomedicine 88 (2007) 52–61. doi:`10.1016/j.cmpb.2007.07.002`.

[22] N. Meshkat, C. E.-Z. Kuo, J. DiStefano, On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and COMBOS: A novel web implementation, PLoS ONE 9 (2014) e110261. doi:`10.1371/journal.pone.0110261`.

[23] K. Thomaseth, M. P. Saccomani, Local identifiability analysis of nonlinear ODE models: How to determine all candidate solutions, IFAC-PapersOnLine 51 (2018) 529–534. doi:`10.1016/j.ifacol.2018.03.089`.

[24] A. Varghese, S. Narasimhan, N. Bhatt, A priori parameter identifiability in complex reaction networks, IFAC-PapersOnLine 51 (2018) 760–765. doi:`10.1016/j.ifacol.2018.09.162`.

[25] M. P. Saccomani, G. Bellu, S. Audoly, L. D'Angió, A new version of DAISY to test structural identifiability of biological models, in: Computational Methods in Systems Biology, 2019, pp. 329–334. doi:`10.1007/978-3-030-31304-3_21`.

[26] H. Hong, A. Ovchinnikov, G. Pogudin, C. Yap, SIAN: Software for structural identifiability analysis of ODE models, Bioinformatics 35 (2019) 2873–2874. doi:`10.1093/bioinformatics/bty1069`.

[27] H. Hong, A. Ovchinnikov, G. Pogudin, C. Yap, Global identifiability of differential models, Communications on Pure and Applied Mathematics 73 (2020) 1831–1879. doi:`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] A. F. Villaverde, Observability and structural identifiability of nonlinear biological systems, Complexity 2019 (2019) 1–12. doi:`10.1155/2019/8497093`.

[29] D. J. Bates, J. D. Hauenstein, N. Meshkat, Identifiability and numerical algebraic geometry, PLOS ONE 14 (2019) e0226299. doi:`10.1371/journal.pone.0226299`.

[30] O. T. Chis, J. R. Banga, E. Balsa-Canto, Structural identifiability of systems biology models: A critical comparison of methods, PLoS ONE 6 (2011) e27755. doi:`10.1371/journal.pone.0027755`.

[31] A. Raue, J. Karlsson, M. P. Saccomani, M. Jirstrand, J. Timmer, Comparison of approaches for parameter identifiability analysis of biological systems, Bioinformatics 30 (2014) 1440–1448. doi:10.1093/bioinformatics/btu006.

[32] D. Joubert, J. D. Stigter, J. Molenaar, An efficient procedure to assist in the re-parametrization of structurally unidentifiable models, Mathematical Biosciences 323 (2020) 108328. doi: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 non-identifiable parameters.

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

[34] S. A. Murphy, A. W. van der Vaart, On profile likelihood, Journal of the American Statistical Association 95 (2000) 449. doi:10.2307/2669386.

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

[36] O. M. Brastein, B. Lie, R. Sharma, N.-O. Skeie, Parameter estimation for externally simulated thermal network models, Energy and Buildings 191 (2019) 200–210. doi: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] C. Kreutz, An easy and efficient approach for testing identifiability, Bioinformatics 34 (2018) 1913–1921. doi: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] A. Raue, C. Kreutz, F. J. Theis, J. Timmer, Joining forces of Bayesian and frequentist methodology: A study for inference in the presence of non-identifiability, Phil. Trans. Roy. Soc. A 371 (2013) 20110544. doi:10.1098/rsta.2011.0544.

[39] S. Gupta, R. E. C. Lee, J. R. Faeder, Parallel tempering with LASSO for model reduction in systems biology, PLOS Computational Biology 16 (2020) e1007669. doi:10.1371/journal.pcbi.1007669.

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

[41] I. David, A. Ricard, A unified model for inclusive inheritance in livestock species, Genetics 212 (2019) 1075–1099. doi:10.1534/genetics.119.302375.

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

[43] W. Zhou, R. Huang, K. Liu, W. Zhang, A novel interval-based approach for quantifying practical parameter identifiability of a lithium-ion battery model, International Journal of Energy Research 44 (2020) 3558–3573. doi:10.1002/er.5118.

[44] K. E. Johnson, G. R. Howard, D. Morgan, E. A. Brenner, A. L. Gardner, R. E. Durrett, W. Mo, A. Al'Khafaji, E. D. Sontag, A. M. Jarrett, T. E. Yankeelov, A. Brock, Integrating transcriptomics and bulk time course data into a mathematical framework to describe and predict therapeutic resistance in cancer, Physical Biology 18 (2020) 016001. doi:10.1088/1478-3975/abb09c.

   *The authors emphasise 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] M. Nihtilä, J. Virkkunen, Practical identifiability of growth and substrate consumption models, Biotechnology and Bioengineering 19 (1977) 1831–1850. doi:10.1002/bit.260191208.

[46] A. Holmberg, On the practical identifiability of microbial growth models incorporating Michaelis-Menten type nonlinearities, Mathematical Biosciences 62 (1982) 23–43. doi:`10.1016/0025-5564(82)90061-X`.

[47] H. Miao, X. Xia, A. S. Perelson, H. Wu, On identifiability of nonlinear ODE models and Aapplications in viral dynamics, SIAM Review 53 (2011) 3–39. doi:`10.1137/090757009`.

[48] C. Gontier, J.-P. Pfister, Identifiability of a binomial synapse, Frontiers in Computational Neuroscience 14 (2020). doi:`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] M. P. Saccomani, K. Thomaseth, The union between structural and practical identifiability makes strength in reducing oncological model complexity: A case study, Complexity 2018 (2018) 1–10. doi:`10.1155/2018/2380650`.

[50] M. C. Neale, M. B. Miller, The use of likelihood-based confidence intervals in genetic models, Behavior Genetics 27 (1997) 113–120. doi:`10.1023/A:1025681223921`.

[51] R. Ochoa-Fernandez, N. B. Abel, F.-G. Wieland, J. Schlegel, L.-A. Koch, J. B. Miller, R. Engesser, G. Giuriani, S. M. Brandl, J. Timmer, W. Weber, T. Ott, R. Simon, M. D. Zurbriggen, Optogenetic control of gene expression in plants in the presence of ambient white light, Nature Methods 17 (2020) 717–725. doi:`10.1038/s41592-020-0868-y`.

[52] N. Schneider, F.-G. Wieland, D. Kong, A. A. M. Fischer, M. Hörner, J. Timmer, H. Ye, W. Weber, Liquid-liquid phase separation of light-inducible transcription factors increases transcription activation in mammalian cells and mice, Science Advances 7 (2021) eabd3568. doi:`10.1126/sciadv.abd3568`.

[53] I. Siekmann, J. Sneyd, E. J. Crampin, MCMC can detect nonidentifiable models, Biophysical Journal 103 (2012) 2275–2286. doi:`10.1016/j.bpj.2012.10.024`.

[54] K. E. Hines, T. R. Middendorf, R. W. Aldrich, Determination of parameter identifiability in nonlinear biophysical models: A Bayesian approach, Journal of General Physiology 143 (2014) 401–416. doi:`10.1085/jgp.201311116`.

[55] Y. Zuo, S. Sarkar, M. T. Corwin, K. Olson, R. D. Badawi, G. Wang, Structural and practical identifiability of dual-input kinetic modeling in dynamic PET of liver inflammation, Physics in Medicine & Biology 64 (2019) 175023. doi:`10.1088/1361-6560/ab1f29`.

[56] M. J. Simpson, R. E. Baker, S. T. Vittadello, O. J. Maclaren, Practical parameter identifiability for spatio-temporal models of cell invasion, Journal of The Royal Society Interface 17 (2020) 20200055. doi:`10.1098/rsif.2020.0055`.

> *By means of spatio-temporal 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] M. Joshi, A. Seidel-Morgenstern, A. Kremling, Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems, Metabolic Engineering 8 (2006) 447–455. doi:`10.1016/j.ymben.2006.04.003`.

[58] L. Sachs, Applied Statistics, Springer Series in Statistics, Springer New York, New York, NY, 1984. doi:`10.1007/978-1-4612-5246-7`.

[59] C. Kreutz, A. Raue, J. Timmer, Likelihood based observability analysis and confidence intervals for predictions of dynamic models, BMC Systems Biology 6 (2012) 120. doi:`10.1186/1752-0509-6-120`.

[60] S. S. Bhonsale, B. Stokbroekx, J. Van Impe, Assessment of the parameter identifiability of population balance models for air jet mills, Computers & Chemical Engineering 143 (2020) 107056. doi:`10.1016/j.compchemeng.2020.107056`.

[61] H. G. Frøysa, H. J. Skaug, G. Alendal, Experimental design for parameter estimation in steady-state linear models of metabolic networks, Mathematical Biosciences 319 (2020) 108291. doi:`10.1016/j.mbs.2019.108291`.

[62] R. Muñoz-Tamayo, L. Puillet, J. B. Daniel, D. Sauvant, O. Martin, M. Taghipoor, P. Blavy, Review: To be or not to be an identifiable model. Is this a relevant question in animal science modelling?, Animal 12 (2018) 701–712. doi:`10.1017/S1751731117002774`.

[63] T. Gross, N. Blüthgen, Identifiability and experimental design in perturbation studies, Bioinformatics 36 (2020) i482–i489. doi:`10.1093/bioinformatics/btaa404`.

[64] L. Bandiera, D. Gomez-Cabeza, J. Gilman, E. Balsa-Canto, F. Menolascina, Optimally designed model selection for synthetic biology, ACS Synthetic Biology 9 (2020) 3134–3144. doi:10.1021/acssynbio.0c00393.

[65] A. Raue, V. Becker, U. Klingmüller, J. Timmer, Identifiability and observability analysis for experimental design in nonlinear dynamical models, Chaos 20 (2010) 45105. doi:10.1063/1.3528102.

[66] B. Steiert, A. Raue, J. Timmer, C. Kreutz, Experimental design for parameter estimation of gene regulatory networks, PLoS ONE 7 (2012) e40052. doi:10.1371/journal.pone.0040052.

[67] A. Tivay, X. Jin, A. K.-Y. Lo, C. G. Scully, J.-O. Hahn, Practical use of regularization in individualizing a mathematical model of cardiovascular hemodynamics using scarce data, Frontiers in Physiology 11 (2020). doi:10.3389/fphys.2020.00452.

[68] Q. P. He, J. Wang, Application of systems engineering principles and techniques in biological big data analytics: A review, Processes 8 (2020) 951. doi:10.3390/pr8080951.

[69] Z. Zi, Sensitivity analysis approaches applied to systems biology models, IET Systems Biology 5 (2011) 336–346. doi:10.1049/iet-syb.2011.0015.

[70] A. Pironet, P. D. Docherty, P. C. Dauby, J. G. Chase, T. Desaive, Practical identifiability analysis of a minimal cardiovascular system model, Computer Methods and Programs in Biomedicine 171 (2019) 53–65. doi:10.1016/j.cmpb.2017.01.005.

[71] T. Maiwald, H. Hass, B. Steiert, J. Vanlier, R. Engesser, A. Raue, F. Kipkeew, H. H. Bock, D. Kaschek, C. Kreutz, J. Timmer, Driving the model to its limit: Profile likelihood based model reduction, PLoS ONE 11 (2016) e0162366. doi:10.1371/journal.pone.0162366.

[72] C. Tönsing, J. Timmer, C. Kreutz, Profile likelihood-based analyses of infectious disease models, Statistical Methods in Medical Research 27 (2018) 1979–1998. doi: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.

[73] L. Cucurull-Sanchez, M. J. Chappell, V. Chelliah, S. Y. Amy Cheung, G. Derks, M. Penney, A. Phipps, R. S. Malik-Sheriff, J. Timmis, M. J. Tindall, P. H. Graaf, P. Vicini, J. W. T. Yates, 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 & Systems Pharmacology 8 (2019) 259–272. doi:10.1002/psp4.12381.

[74] D. L. Janzén, M. Jirstrand, M. J. Chappell, N. D. Evans, Three novel approaches to structural identifiability analysis in mixed-effects models, Computer Methods and Programs in Biomedicine 171 (2019) 141–152. doi:10.1016/j.cmpb.2016.04.024.