Testing structural identifiability by a simple scaling method

Mario Castro[a,b,*], Rob J. de Boer[c]

a Grupo Interdisciplinar de Sistemas Complejos (GISC)
b Instituto de Investigación Tecnológica (IIT), Universidad Pontificia Comillas, Madrid, E28015, Spain.
c Theoretical Biology and Bioinformatics, Utrecht University, Utrecht, The Netherlands.

‡All authors also contributed equally to this work.
* marioc@comillas.edu

Abstract

Successful mathematical modeling of biological processes relies on the expertise of the modeler to capture the essential mechanisms in the process at hand and on the ability to extract useful information from empirical data. The very structure of the model limits the ability to infer numerical values for the parameters, a concept referred to as structural identifiability. Most of the available methods to test the structural identifiability of a model are either too complex mathematically for the general practitioner to be applied, or require involved calculations or numerical computation for complex non-linear models. In this work, we present a new analytical method to test structural identifiability of models based on ordinary differential equations, based on the invariance of the equations under the scaling transformation of its parameters. The method is based on rigorous mathematical results but it is easy and quick to apply, even to test the identifiability of sophisticated highly non-linear models. We illustrate our method by example and compare its performance with other existing methods in the literature.

Author summary

Theoretical Biology is a useful approach to explain, generate hypotheses, or discriminate among competing theories. A well-formulated model has to be complex enough to capture the relevant mechanisms of the problem, and simple enough to be fitted to data. Structural identifiability tests aim to recognize, in advance, if the structure of the model allows parameter fitting even with unlimited high-quality data. Available methods require advanced mathematical skills, or are too costly for high-dimensional non-linear models. We propose an analytical method based on scale invariance of the equations. It provides definite answers to the structural identifiability problem while being simple enough to be performed in a few lines of calculations without any computational aid. It favorably compares with other existing methods.

Introduction

Mathematical models contribute to our understanding of Biology in several ways ranging from the quantification of biological processes to reconciling conflicting experiments [1]. In many cases, this requires formulating a mathematical model and extracting quantitative estimates of its parameters from the experimental data. Parameters are typically unknown constants that change the behavior of the model. While it is usually recognized that parameter estimation requires the availability of sufficient informative data, sometimes it is not possible to estimate all parameters due to the structure of the model (whatever the quantity or quality of the data). This inability is referred to as 'structural identifiability', a concept introduced decades ago by Bellman and Åström [2,3], as opposed to the 'practical identifiability' that depends on limitations set by the data. Practical identifiability has important consequences that can lead to questionable interpretations of the data leading to some recent controversy around this point [4,5]. Structural identifiability is a necessary condition for model fitting and should be used before any attempt to extract information about the parameters, and as a test of the applicability of the model itself. Structural

identifiability can be qualified as global or local. Global structural identifiability tests the ability to estimate unique sets of parameters, while local (or simply, structural identifiability) means that parameters can be estimated only in a limited subset of the space of parameters, *i.e.*, only combinations of parameters are identifiable [6–10]. In practical terms, these definitions can be translated into the language of sensitivity analysis as identifiability requires that (i) the columns of the sensitivity (or, equivalently, the elasticity) matrix are linearly independent, and (ii) each of its columns has at least one large entry [11, 12].
Traditionally, work primarily focused on linear systems [2, 3, 13] based on ordinary differential equations (ODE). For non-linear models, those methods cannot be applied, so many methods have been proposed in the literature to address structural identifiability. Early attempts were based on power series expansions of the original non-linear system [14], the similarity transformation method [15–17] or the so-called direct-test method proposed by Denis-Vidal and Joly-Blanchard [18, 19]. These methods exploit the definition of identifiability either analytically [18] or numerically [20–25], but they are not generically suitable for high-dimensional problems. Xia and Moog [6, 26] proposed an alternative to these classical methods based on the implicit function theorem, but this method also becomes involved to apply for complex models [27].
Another approach that is becoming mainstream is based on the framework of differential algebra [28–31]. These methods are also difficult to apply, requiring advanced mathematical skills and, in some cases, replace highly non-linear terms by polynomial approximations that simplify the analysis. On the positive side, they are based on rigorous mathematical theories, are suitable for non-linear models and, more importantly, they can be coded using existing symbolic computational libraries. In this regard, it is worth mentioning DAISY [32], GenSSI [33], COMBOS [34] or, more recently, SIAN [35].
In almost all cases, the major disadvantage of these methods is their difficulty to apply them to even a few differential equations, hence requiring advanced mathematical skills and/or dedicated numerical or symbolic software (that is frequently unable to handle the complexity of the problem). This explains why, despite the huge volume of publications in the field of theoretical biology, only a few address parameter identifiability explicitly. In this paper, we introduce a simple method to assess local structural identifiability of ODE models that reduces the complexity of existing methods and can bring identifiability testing to a broader audience. Our method is based on simple scaling transformations, and the solution of simple sparse systems of equations. Identifiability for stochastic models [36] is out of the scope of our work.

Method

A couple of motivating examples

Consider a simple *death* model in which the death rate is the product of two parameters λ_1 and λ_2, namely

$$\frac{dx}{dt} = -\lambda_1\lambda_2 x, \qquad x(0) = x_0, \tag{1}$$

with the solution

$$x = x(\lambda_1, \lambda_2, t) = x_0 e^{-\lambda_1\lambda_2 t}. \tag{2}$$

It is evident that from an experiment only the product $\lambda_1\lambda_2$ can be inferred, and not any of the two independently. Following the 'actionable' definition in Ref. [11], local structural identifiability is directly linked to the linear independence of the columns of the elasticity (or, sensitivity) matrix. In this case, the elasticity matrix would be simply a 1×2 matrix,

$$K = (K_1\ K_2)\,,$$

with

$$K_1 \equiv \frac{\lambda_1}{x}\frac{\partial x}{\partial \lambda_1}\ \text{, and} \quad K_2 \equiv \frac{\lambda_2}{x}\frac{\partial x}{\partial \lambda_2}\ . \tag{3}$$

We now propose to multiply λ_1 with a generic scale factor u, and to divide λ_2 by the same factor, such that the solution remains invariant. Deriving the scaled solution of eq. (2) with respect to that scale factor u, and by the chain rule,

$$\frac{dx}{du} = 0 \text{ (as } u \text{ is arbitrary)} \tag{4}$$

and, also,

$$\frac{dx(u\lambda_1, \lambda_2/u, t)}{du} = \frac{\partial x}{\partial \lambda_1}\lambda_1 - \frac{\lambda_2}{u^2}\frac{\partial x}{\partial \lambda_2} = 0 \tag{5}$$

where the last equality follows from Eq. (4)
Rearranging Eq. (5) and dividing by x,

$$\frac{\lambda_1}{x}\frac{\partial x}{\partial \lambda_1} = \frac{\lambda_2}{u^2 x}\frac{\partial x}{\partial \lambda_2} \Rightarrow K_1 = \frac{1}{u^2}K_2, \tag{6}$$

so both *columns* of the elasticity matrix are linearly dependent and, accordingly, λ_1 and λ_2 are unidentifiable. In this case we had complete knowledge of the solution, and consequently, it was straightforward to find the right way to introduce the scaling u. Fortunately, this simple scaling calculation can also be performed directly on eq. (1). Introducing two unknown scaling factors, u_1 and u_2, into that equation,

$$\frac{dx}{dt} = -u_1\lambda_1 u_2\lambda_2 x \ .$$

Requiring that this remains identical (or, more formally, *invariant*) to Eq. (1), i.e., $\lambda_1\lambda_2 x = u_1\lambda_1 u_2\lambda_2 x$, hence $u_1 u_2 = 1$. The fact that u_1 and u_2 cannot be solved individually, also means that the real values of λ_1 and λ_2 cannot be determined, namely both parameters are unidentifiable.
Next consider a death model with immigration:

$$\frac{dx}{dt} = \lambda_1 - \lambda_2 x. \tag{7}$$

In this case, to leave the system invariant we need to find u_1 and u_2 such that

$$\lambda_1 - \lambda_2 x(t) = u_1\lambda_1 - u_2\lambda_2 x(t)$$

for all values of x at any time. Rearranging the latter equation,

$$(1-u_1)\lambda_1 = (1-u_2)\lambda_2 x(t),$$

where the left-hand side of the last equation is a constant and the right-hand side depends on time. Hence the only possible solution to the latter equation is $u_1 = u_2 = 1$ implying that both λ_1 and λ_2 are locally identifiabile. Notice the difference with the preceding case, Eq. (1), in which an infinite number of combinations of the scaling factors satisfy the invariance condition.
These simple examples illustrate how scaling invariance of the model equations can be used to determine whether the parameters are unidentifiable or not. We prove this result more rigorously in the Supporting Information.

Description of the method

Let us define a general ODE model characterized by the time evolution of n variables, $x_i(t)$, depending on m parameters λ_j,

$$\frac{dx_i}{dt} = f_i(x_1,\ldots,x_r,x_{r+1}\ldots x_n;\lambda_1,\ldots,\lambda_m) \qquad i=1,\ldots,n \ , \tag{8a}$$

$$x_i(0) = x_{i,0}, \qquad i=1,\ldots,n \ , \tag{8b}$$

where the functions f_i depend on the specific details of the problem at hand and $x_{i,0}$ are the initial conditions. We need to distinguish between those variables that can be observed (measured) in the experiment, $x_1\ldots x_r$, and those which cannot (they are often referred to as *latent* variables), $x_{r+1}\ldots x_n$. According to Sec. 1 in the Supporting Information, each function f_i is split into M functional independent *summands*, f_{ik},

$$f_i(x_1,\ldots,x_r,x_{r+1}\ldots x_n;\lambda_1,\ldots,\lambda_m) = \sum_{k=1}^{M} f_{ik}(\tilde{x}_k,\tilde{\lambda}_k),$$

having the property that f_{ik} is functionally independent of f_{il} for every $k \neq l$. For brevity, $\tilde{x}_k, \tilde{\lambda}_k$ denote the subset of variables and parameters included in the function f_{ik}. Typical examples of functionally independent functions are $f_{11} = ax_1$, $f_{12} = bx_1x_3$, $f_{13} = (c+x_4)^{-1}$, and examples of dependent functions would be

Box 1: Summary of the scale invariance local structural identifiability method introduced in this work.

1. Scale all parameters and all unobserved variables by unknown scaling factors, u:

$$\begin{aligned} \lambda_i &\to u_{\lambda_i}\lambda_i \quad i = 1, \dots, m \\ x_j &\to u_{x_j} x_j \quad j = r+1, \dots, n \end{aligned}$$

   and substitute them into Eqs. (10) below.

2. Equate each functionally independent function, f_{ik}, to its scaled version. Namely,

$$f_{ik}(\tilde{x}, \tilde{\lambda}) = \frac{1}{u_{x_i}} f_{ik}(u_{\tilde{x}}\tilde{x}, u_{\tilde{\lambda}}\tilde{\lambda}) \tag{9}$$

   where $u_{x_i} = 1$ for $1 \leq i \leq r$ and the prefactor in the right-hand side of the equation comes from the scaling of $\frac{dx_i}{dt} \to u_{x_i}\frac{dx_i}{dt}$. From Eq. (8b) it follows that $u_{x_i} = u_{x_{i,0}}$.

3. From the Eqs. (9), find combinations of the scaling factors u that leave the system invariant. Hereafter, we will denote these as the **identifiability equations** of the model.

4. Only the parameters λ_i with a solution $u_i = 1$ are identifiable.

$f_{11} = ax_1x_2$ and $f_{12} = bx_1x_2$ (see Supporting Information for details). Note that it is not required that f_{ij} and f_{kj} are independent (as they appear in different equations). For instance, in the example in Eq. (7)

$$f_{11} = \lambda_1 \text{ , and } \quad f_{12} = -\lambda_2 x \ .$$

We summarize our method in Box 1.

In summary, our method reduces the complexity of finding identifiable parameters to finding which scaling factors do not satisfy the trivial solution $u_i = 1$. In the literature, when a scaling factor is related to one of the latent variables $x_{r+1} \dots x_n$, if $u_{x_k} = 1$, then x_k is said to be *observable* [10]. Thus, our method addresses at the same time identifiability and observability. Additionally, irreducible equations involving two or more parameters provide the so-called identifiable groups of variables that cannot be fitted independently. In the case of the pure death model above, the identifiability equation $u_{\lambda_1}u_{\lambda_2} = 1$ is a signature of the unidentifiable group $\lambda_1\lambda_2$. This is interesting as groups involving latent variables (for instance, $u_{x_j}u_{\lambda_k}$) would inform future experiments aimed to observe that variable and decouple that group.

It is also worth mentioning that our identifiability test (illustrated by example in the Supporting Information) provides a simple way to find a type of symmetry that is related to scale invariance. More sophisticated methods have been introduced in the literature to address other symmetries [37–39] using the theory of Lie group transformations, however, that approach involves complex calculations assisted by symbolic computations.

Results

The main result

Consider a model described by a set of n ordinary differential equations (ODE)

$$\frac{dx_i}{dt} = f_i(x_1, \dots, x_r, x_{r+1} \dots x_n; \lambda_1, \dots, \lambda_m) = \sum_{k=1}^{M} f_{ik}(\tilde{x}_k, \tilde{\lambda}_k), \tag{10}$$

where f_{ik} is functionally independent of f_{il} for every $k \neq l$ (namely, they satisfy the generalized Wronskian theorem; see the Supporting Information). For the sake of simplicity, we denote $\tilde{x}_k$ and $\tilde{\lambda}_k$ the subset of variables and parameters of function f_{ik}.

Motivated by Eqs. (1)-(3), we seek for scaling of the parameters that leave the system invariant. As we prove below, this invariance (or lack of) is related to the identifiability of the parameters. Hence, if we define the following scaling transformation:

$$\lambda_i \to u_{\lambda_i}\lambda_i, \quad i = 1, \ldots, m \qquad x_j \to u_{x_j}x_j, \quad j = r+1, \ldots, n \tag{11}$$

(where the variables $x_1 \ldots x_r$ are unmodified as we can measure them in the experiment) we can write the following set of re-scaled equations:

$$\frac{dx_i}{dt} = \sum_{k=1}^{M} f_{ik}(u_{\tilde{x}_k}\tilde{x}_k, u_{\tilde{\lambda}_k}\tilde{\lambda}_k),\ i = 1, \ldots, r \tag{12a}$$

$$x_i = x_{i,0},\ i = 1, \ldots, r \tag{12b}$$

$$u_{x_i}\frac{dx_i}{dt} = \sum_{k=1}^{M} f_{ik}(u_{\tilde{x}_k}\tilde{x}_k, u_{\tilde{\lambda}_k}\tilde{\lambda}_k),\ i = r+1, \ldots, n \tag{12c}$$

$$u_{x_i}x_i = u_{x_{i,0}}x_{i,0},\ i = r+1, \ldots, n \tag{12d}$$

where M is the number of functional independent summands in the equation. It is convenient to rewrite Eq. (12c) as

$$\frac{dx_i}{dt} = \frac{1}{u_{x_i}}\sum_{k=1}^{M} f_{ik}(u_{\tilde{x}_k}\tilde{x}_k, u_{\tilde{\lambda}_k}\tilde{\lambda}_k),\ i = r+1, \ldots n \tag{13}$$

to perform the scale invariance analysis below in a simpler way.
If the solution is invariant under this transformation, then the right-hand sides of Eq. (10) and, consequently Eqs (12) should be equal. Besides, by the functional linear independence of the functions f_{ik} we can split each summand. Thus,

$$f_{ik}(\tilde{x}_k, \tilde{\lambda}_k), = f_{ik}(u_{\tilde{x}_k}\tilde{x}_k, u_{\tilde{\lambda}_k}\tilde{\lambda}_k), \quad i = 1, \ldots, r \tag{14}$$

and

$$f_{ik}(\tilde{x}_k, \tilde{\lambda}_k), = \frac{1}{u_{x_i}} f_{ik}(u_{\tilde{x}_k}\tilde{x}_k, u_{\tilde{\lambda}_k}\tilde{\lambda}_k), \quad i = r+1, \ldots, n \tag{15}$$

These new set of equations are much easier to solve than the ones that we would obtain from Eqs (12a)-(12c) (which would be equivalent to the so-called direct-test method [18]).
We can express the solution of these equations as

$$u_{\lambda_k} = F(u_{x_{m_1}}, u_{x_{m_2}}, \ldots, u_{\lambda_{j_1}}, u_{\lambda_{j_2}}, \ldots), \tag{16}$$

We denote these the **identifiability equations** of the model. For each parameter k, the identifiability equation will depend only on a few other scaling factors $m_1, m_2, \ldots$.
If take the partial derivative of the transformed solution

$$x_i(x_1, \ldots, x_r, u_{x_{r+1}}x_{r+1} \ldots u_{x_n}x_n; u_{\lambda_1}\lambda_1, \ldots, u_{\lambda_m}\lambda_m)$$

with respect to u_{λ_k}, we find (by the chain rule)

$$\frac{\partial x_i}{\partial u_{\lambda_k}} = \frac{\partial x_i}{\partial \lambda_k}\lambda_k + \frac{\partial x_i}{\partial \lambda_{j_1}}\lambda_{j_1}\frac{\partial u_{\lambda_k}}{\partial u_{\lambda_{j_1}}} + \frac{\partial x_i}{\partial \lambda_{j_2}}\lambda_{j_2}\frac{\partial u_{\lambda_k}}{\partial u_{\lambda_{j_2}}} + \ldots \tag{17}$$

On the other hand, as the solution is invariant with respect to the scaling transformation, x_i should not depend on the arbitrary scaling factors, thus

$$\frac{\partial x_i}{\partial u_{\lambda_k}} = 0.$$

Finally, dividing Eq. (17) by x_i and defining

$$\beta_{kj} \equiv \frac{\partial u_{\lambda_k}}{\partial u_{\lambda_j}}$$

and the elasticity matrix K with elements K_{ij} given by

$$K_{ij}(x_1, \ldots, x_r, x_{r+1} \ldots x_n; \lambda_1, \ldots, \lambda_m) = \frac{\lambda_j}{x_i} \frac{\partial x_i}{\partial \lambda_j} \tag{18}$$

we find

$$K_{ik} + \beta_{kj_1} K_{ij_1} + \beta_{kj_2} K_{ij_2} + \ldots = 0,$$

At this point we have two possibilities. Either all the coefficients $\beta_{kj} = 0$, meaning that Eq. (16) is not satisfied (we will simply have $u_{\lambda_k} = 1$), or the columns of the elasticity matrix related to the parameter λ_k has linearly dependent columns, and then the parameters λ_k, $\lambda_{j_1}, \lambda_{j_2}, \ldots$ are unidentifiable.
On the other hand, if the solution of the identifiability equation is $u_{\lambda_k} = 1$ for some parameter, then the column of the elasticity matrix coming from that parameter is linearly independent of the others, and the parameter is **local** structurally identifiable. The adjective "local" is required because the method stems on the continuity of the derivative of the solution of Eq. (17). Thus, it is unable to capture any discrete (*countable*) transformations like, for instance, those related to exchanging two parameters of the model.
Finally, by Eq. (12d) we find that $u_{x_i} = u_{x_{i,0}}$, so an initial condition is identifiable whenever its corresponding variable is observable.

Comparison with other methods

We have applied the method outlined in Box 1 to 13 different models defined and analyzed in detail in the Supporting Information. The choice is based on two criteria: on the one hand, models 1-5 are included for *pedagogical* purposes. They are simple enough to illustrate the method and most of the existing methods also provide definite answers. Models 6-13 were chosen because they have previously been analyzed using the methods summarized in the Introduction and in Table 1. This allows us to put our method in direct competition with those methods and to highlight their merits and limitations.
The results of this comparison are summarized in Table 2, which is inspired by a similar table in Ref. [7].

Table 1. List of current methods testing structural identifiability. We introduce here the acronyms referred to in Table 2.

Method	Acronym	Main Ref.	Pros	Cons
Direct test method	DT	[18, 20]	Simple	Limited
Implicit function theorem	IFT	[26]	Software	Limited
Taylor series approach	TS	[14]	Simple	Computationally Expensive
Generating series approach	GS	[13]	Simple, Software	Computationally Expensive
Similarity Transformation	ST	[16]	Software	Computationally Expensive
Differential algebra	DA	[29, 32, 34]	Software, Conclusive	Limited, Comp. Expensive
Reaction Network theory	RNT	[40, 41]	Simple, Hybrid with other	Only reaction systems
STRIKE-GOLDD	SG	[9, 22]	Powerful, Software	Computationally Expensive
Scaling Invariance Method	**SIM**	This work	Simple, Widely applicable	Only Local Identifiability

Discussion and Conclusions

Table 2 shows that our method can handle any complex model and provides a local structural identifiability criterion that is compatible with those methods capable of producing an answer. Thus, our method is widely applicable. It is worth noting that in several cases where our scaling method comes with a conclusive answer, other more complicated methods cannot address those cases (rightmost column in the table). As any global structural identifiable model is also local, our results are compatible with those methods that can address that difference.
It is worth emphasizing that we have performed our test by hand, as illustrated in the Supporting Information, and that, after some practice (and using some interesting *motifs* as having sums of different parameters, or the coefficients related to *diagonal* terms in the system of equations) the calculations can be

Table 2. Summary of models compared in the literature: The number in brackets in the Model Name column correspond to the number of observed variables. Model Numbers correspond to those in Table 1 in the Supporting Information. The acronyms for the methods are summarized in Table 1. This table has been inspired by Table 1 in Ref. [7].

Model name	**Main Ref.**	**Model Number**	**Global Struct. Id.**	**Local Struct. Id.**	**Unidentifiable**	**Not Conclusive Not Applicable**
Goodwin model (1)	[7]	6			SG,**SIM**	TS,GS,ST,DT,DA,IFT,RNT
Goodwin model (all)	[7]	6bis		TS,GS,IFT,RNT	DA,SG,**SIM**	ST,DT
Circadian clock model	[42]	7			TS,GS,RNT,SG,**SIM**	ST,DT,DA,IFT
HIV model (1)	[6,43]	8			All	
HIV model (2)	[6,43]	8bis	DA,IFT,RNT	TS,GS,**SIM**		DT,ST
Linear HIV model (1)	[6,43,44]	8ter	DA,IFT,RNT,SG	DT,ST,TS,GS,**SIM**		
Glycolysis model	[45]	9	GS,DA,RNT	TS,RNT,**SIM**		ST,DT
High dimensional model	[42]	10	TS,GS,DA,RNT	IFT,**SIM**		ST,DT
NF-κ model B (1)	[46]	11			SG, **SIM**	TS,GS,ST,DT,DA,IFT,RNT
NF-κ model B (2)	[46]	11bis	GS,RNT	TS,**SIM**	SG	ST,DT,DA,IFT
Pharmacokinetics model (1)	[47]	12		TS,GS,RNT,SG,**SIM**		ST,DT,DA,IFT
Pharmacokinetics model (2)	[47]	12bis	DA	GS,SG,**SIM**		ST,DT,IFT,RNT
Within-host virus model	[27]	13	DA	**SIM**		TS,GS,ST,DT,IFT,RNT

made in a few minutes. This contrasts with the most sophisticated methods that, by hand, can fill several pages [27] or take hours using symbolic computation packages.

Together, this broad applicable and simplicity are the main features of our method and this may attract the interest of mathematical modelers and spread the *culture* of checking structural identifiability as a mandatory step when fitting experimental data.

We would like to highlight a connection with the so-called Buckingham-Π theorem of dimensional analysis [48]. In some sense, the scale invariance property is related to the principle of dimensional homogeneity, *i.e.*, the constraints on the functional form of the independent variables with the parameters. Our identifiability equations are therefore similar to finding the so-called Π-groups in the theorem.

A limitation of the method is that it is restricted to testing local identifiability. This is implicit in the differentiability of the elasticity matrix which, by definition, is a local operation. Discrete symmetries are not captured, and more sophisticated methods (based on Lie group transformations [39]) are required. However, simple manipulation of the equations to remove the latent variables can improve the explanatory power of the method and might capture those discrete symmetries (see Sec. 3.8 of the Supporting Information). We leave that extension for future developments.

Finally, in this work we have chosen to solve the scaling factor equations directly as it is easy to perform with pen and paper. However, if we were to redefine the scaling factors as $u_i = e^{w_i}$, the new factors w_i would obey a linear system of homogeneous equations. It is therefore expected that the problem of identifiability is related to the rank of the matrix defining the linear system of equations. In that regard, the theorems presented in the Supporting Information could be supplemented with generic results on homogeneous systems of equations. Thus, our results provide a solid ground for the method and indicate a venue for further development in other systems like delay-differential or partial differential equations. Finally, while we emphasize the simplicity of the method, it is obviously amenable to be implemented using symbolic computation packages.

Supporting information

In the Supporting Information we collect the theorems sustaining the method and a catalogue of models with a detailed computation of the identifiability equations that were used to build Table 2.

Acknowledgments

This work has been partially supported by grant FIS2016-78883-C2- 2-P (Ministerio de Economia, Industria y Competitividad - Agencia Estatal de Investigacion). This work was initiated during summer visits of the authors to the Los Alamos National Laboratory, and we thank Nick Hengartner and Alan Perelson (LANL) for their hospitality and helpful comments on this work, and the Santa Fe Institute for supporting the summer visits of RdB.

References

1. Castro M, Lythe G, Molina-París C, Ribeiro RM. Mathematics in Modern Immunology. Interface focus. 2016;6(2):20150093.
2. Bellman R, Åström KJ. On structural identifiability. Mathematical Biosciences. 1970;7(3-4):329–339.
3. Jacquez JA, et al. Compartmental analysis in biology and medicine. New York, Elsevier Pub. Co.; 1972.
4. Balsa-Canto E, Alonso-del Real J, Querol A. Mixed growth curve data do not suffice to fully characterize the dynamics of mixed cultures. Proceedings of the National Academy of Sciences. 2020;117(2):811.
5. Ram Y, Obolski U, Feldman MW, Berman J, Hadany L. Reply to Balsa-Canto et al.: Growth models are applicable to growth data, not to stationary-phase data. Proceedings of the National Academy of Sciences. 2020;117(2):814–815.
6. Miao H, Xia X, Perelson AS, Wu H. On the identifiability of nonlinear ODE models and applications in viral dynamics. SIAM review. 2011;53(1):3–39.
7. Chis OT, Banga JR, Balsa-Canto E. Structural identifiability of systems biology models: a critical comparison of methods. PloS one. 2011;6(11):e27755.
8. Villaverde AF, Barreiro A. Identifiability of large nonlinear biochemical networks. Match Commun Math Comput Chem (Mulheim an der Ruhr, Germany). 2016;76(2):259–276.
9. Villaverde AF, Barreiro A, Papachristodoulou A. Structural Identifiability of Dynamic Systems Biology Models. PLoS Computational Biology. 2016;12(10):1–22. doi:10.1371/journal.pcbi.1005153.
10. Villaverde AF, Tsiantis N, Banga JR. Full observability and estimation of unknown inputs, states and parameters of nonlinear biological models. Journal of the Royal Society Interface. 2019;16(156):20190043.
11. Jaqaman K, Danuser G. Linking data to models: data regression. Nature Reviews Molecular Cell Biology. 2006;7(11):813.
12. Komorowski M, Costa MJ, Rand DA, Stumpf MP. Sensitivity, robustness, and identifiability in stochastic chemical kinetics models. Proceedings of the National Academy of Sciences. 2011;108(21):8645–8650.
13. Walter E, Lecourtier Y. Unidentifiable compartmental models: What to do? Mathematical Biosciences. 1981;56(1-2):1–25.
14. Pohjanpalo H. System identifiability based on the power series expansion of the solution. Mathematical Biosciences. 1978;41(1-2):21–33.
15. Vajda S, Rabitz H. State isomorphism approach to global identifiability of nonlinear systems. IEEE Transactions on Automatic Control. 1989;34(2):220–223.
16. Vajda S, Godfrey KR, Rabitz H. Similarity transformation approach to identifiability analysis of nonlinear compartmental models. Mathematical Biosciences. 1989;93(2):217–248.
17. Chappell MJ, Godfrey KR. Structural identifiability of the parameters of a nonlinear batch reactor model. Mathematical Biosciences. 1992;108(2):241–251.
18. Denis-Vidal L, Joly-Blanchard G. An easy to check criterion for (un) indentifiability of uncontrolled systems and its applications. IEEE Transactions on Automatic Control. 2000;45(4):768–771.
19. Raksanyi A, Lecourtier Y, Walter E, Venot A. Identifiability and distinguishability testing via computer algebra. Mathematical Biosciences. 1985;77(1-2):245–266.

20. Walter E, Braems I, Jaulin L, Kieffer M. Guaranteed numerical computation as an alternative to computer algebra for testing models for identifiability. In: Numerical Software with Result Verification. Springer; 2004. p. 124–131.

21. Maiwald T, Hass H, Steiert B, Vanlier J, Engesser R, Raue A, et al. Driving the model to its limit: profile likelihood based model reduction. PloS one. 2016;11(9):e0162366.

22. Villaverde AF, Barreiro A, Papachristodoulou A. Structural Identifiability Analysis via Extended Observability and Decomposition. IFAC-PapersOnLine. 2016;49(26):171–177.

23. Kreutz C. An easy and efficient approach for testing identifiability. Bioinformatics. 2018;34(11):1913–1921.

24. Tönsing C, Timmer J, Kreutz C. Profile likelihood-based analyses of infectious disease models. Statistical methods in medical research. 2018;27(7):1979–1998.

25. Stigter JD, Molenaar J. A fast algorithm to assess local structural identifiability. Automatica. 2015;58:118–124.

26. Xia X, Moog CH. Identifiability of nonlinear systems with application to HIV/AIDS models. IEEE transactions on automatic control. 2003;48(2):330–336.

27. Koelle K, Farrell AP, Brooke CB, Ke R. Within-host infectious disease models accommodating cellular coinfection, with an application to influenza. Virus evolution. 2019;5(2):vez018.

28. Walter E, Pronzato L. On the identifiability and distinguishability of nonlinear parametric models. Mathematics and computers in simulation. 1996;42(2-3):125–134.

29. Ljung L, Glad T. On global identifiability for arbitrary model parametrizations. Automatica. 1994;30(2):265–276.

30. Ollivier F. Identifiabilité et identification: du Calcul Formel au Calcul Numérique? In: ESAIM: Proceedings. vol. 9. EDP Sciences; 2000. p. 93–99.

31. Meshkat N, Eisenberg M, DiStefano III JJ. An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner Bases. Mathematical Biosciences. 2009;222(2):61–72.

32. Bellu G, Saccomani MP, Audoly S, D'Angiò L. DAISY: A new software tool to test global identifiability of biological and physiological systems. Computer methods and programs in biomedicine. 2007;88(1):52–61.

33. Chiş O, Banga JR, Balsa-Canto E. GenSSI: a software toolbox for structural identifiability analysis of biological models. Bioinformatics. 2011;27(18):2610–2611.

34. Meshkat N, Kuo CEz, DiStefano III J. On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and COMBOS: a novel web implementation. PLoS One. 2014;9(10):e110261.

35. Hong H, Ovchinnikov A, Pogudin G, Yap C. SIAN: software for structural identifiability analysis of ODE models. Bioinformatics. 2019;35(16):2873–2874.

36. Brouwer AF, Meza R, Eisenberg MC. Parameter estimation for multistage clonal expansion models from cancer incidence data: A practical identifiability analysis. PLoS computational biology. 2017;13(3):e1005431.

37. Yates JWT, Evans ND, Chappell MJ. Structural identifiability analysis via symmetries of differential equations. Automatica. 2009;45(11):2585–2591. doi:10.1016/j.automatica.2009.07.009.

38. Anguelova M, Karlsson J, Jirstrand M. Minimal output sets for identifiability. Mathematical Biosciences. 2012;239(1):139–153.

39. Merkt B, Timmer J, Kaschek D. Higher-order Lie symmetries in identifiability and predictability analysis of dynamic models. Physical Review E. 2015;92(1):012920.

40. Craciun G, Kim J, Pantea C, Rempala GA. Statistical model for biochemical network inference. Communications in Statistics-Simulation and Computation. 2013;42(1):121–137.

41. Davidescu FP, Jørgensen SB. Structural parameter identifiability analysis for dynamic reaction networks. Chemical Engineering Science. 2008;63(19):4754–4762.

42. Locke J, Millar A, Turner M. Modelling genetic networks with noisy and varied experimental data: the circadian clock in Arabidopsis thaliana. Journal of theoretical biology. 2005;234(3):383–393.

43. Wu H, Zhu H, Miao H, Perelson AS. Parameter identifiability and estimation of HIV/AIDS dynamic models. Bulletin of mathematical biology. 2008;70(3):785–799.

44. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature. 1995;373(6510):123–126.

45. Bartl M, Kötzing M, Kaleta C, Schuster S, Li P. Just-in-time activation of a glycolysis inspired metabolic network-solution with a dynamic optimization approach. In: Crossing Borders within the ABC: Automation, Biomedical Engineering and Computer Science. vol. 55; 2010. p. 217–222.

46. Lipniacki T, Paszek P, Brasier AR, Luxon B, Kimmel M. Mathematical model of NF-κB regulatory module. Journal of theoretical biology. 2004;228(2):195–215.

47. Domurado M, Domurado D, Vansteenkiste S, De Marre A, Schacht E. Glucose oxidase as a tool to study in vivo the interaction of glycosylated polymers with the mannose receptor of macrophages. Journal of controlled release. 1995;33(1):115–123.

48. Buckingham E. Illustrations of the use of dimensional analysis on physically similar systems. Physics Review. 1914;4(4):354–377.