ASAS-NANP symposium: Mathematical Modeling in Animal Nutrition: The power of identifiability analysis for dynamic modeling in animal science:a practitioner approach

Rafael Muñoz-Tamayo, Luis O Tedeschi

To cite this version:

HAL Id: hal-04302072
https://hal.inrae.fr/hal-04302072

Submitted on 23 Nov 2023

OXFORD

AMERICAN SOCIETY OF **ANIMAL SCIENCE**

ASAS-NANP symposium: Mathematical Modeling in Animal Nutrition: The power of identifiability analysis for dynamic modeling in animal science: a practitioner approach

Rafael Muñoz-Tamayo[†,1] and Luis O. Tedeschi[‡]

[†]Université Paris-Saclay, INRAE, AgroParisTech, UMR Modélisation Systémique Appliquée aux Ruminants, 91120 Palaiseau, France
[‡]Department of Animal Science, Texas A&M University, College Station, TX 77843-2471, USA
[1]Corresponding author: rafael.munoz-tamayo@inrae.fr.

Abstract

Constructing dynamic mathematical models of biological systems requires estimating unknown parameters from available experimental data, usually using a statistical fitting procedure. This procedure is usually called parameter identification, parameter estimation, model fitting, or model calibration. In animal science, parameter identification is often performed without analytic considerations on the possibility of determining unique values of the model parameters. These analytical studies are related to the mathematical property of structural identifiability, which refers to the theoretical ability to recover unique values of the model parameters from the measures defined in an experimental setup and use the model structure as the sole basis. The structural identifiability analysis is a powerful tool for model construction because it informs whether the parameter identification problem is well-posed (i.e., the problem has a unique solution). Structural identifiability analysis is helpful to determine which actions (e.g., model reparameterization, choice of new data measurements, and change of the model structure) are needed to render the model parameters identifiable (when possible). The mathematical technicalities associated with structural identifiability analysis are very sophisticated. However, the development of dedicated, freely available software tools enables the application of identifiability analysis without needing to be an expert in mathematics and computer programming. We refer to such a non-expert user as a practitioner for hands-on purposes. However, a practitioner should be familiar with the model construction and software implementation process. In this paper, we propose to adopt a practitioner approach that takes advantage of available software tools to integrate identifiability analysis in the modeling practice in the animal science field. The application of structural identifiability implies switching our regard of the parameter identification problem as a downstream process (after data collection) to an upstream process (before data collection) where experiment design is applied to guarantee identifiability. This upstream approach will substantially improve the workflow of model construction toward robust and valuable models in animal science. Illustrative examples with different levels of complexity support our work. The source codes of the examples were provided for learning purposes and to promote open science practices.

Lay Summary

When modeling biological systems, one major step of the modeling exercise is connecting the theory (the model) with the reality (the data). Such a connection passes through the resolution of the parameter identification (model calibration) problem, which aims at finding a set of parameters that best fits the variables predicted by the model to the data. Traditionally, the parameter identification step is often addressed like a downstream process (after data collection). Using this traditional approach, the modeler has minimal room for maneuvering to improve the model's accuracy. This paper discusses the benefits of adopting an upstream approach (before data collection) during the model construction phase. This approach capitalizes on the identifiability analysis, a powerful tool seldom applied in dynamic models of the animal science domain, likely because of the lack of awareness or the specialized mathematical technicalities involved in the identifiability analysis. In this paper, we illustrate that the modeling community in animal science can easily integrate identifiability analysis in their model developments following a practitioner approach taking advantage of a variety of freely available software tools dedicated to identifiability testing.

Introduction

Modeling the dynamics of a biological system is an exercise of translating the knowledge of the phenomena that drives system behavior into ordinary differential equations (**ODE**). Its state variables (sometimes called compartments in animal science literature) and its parameters define a dynamic mathematical model. The parameter values are often unknown and must be estimated from experimental data via parameter identification (also termed parameter estimation, model calibration, or model fitting). Parameter identification is the mathematical process of finding the numerical values of the model parameters that best fit the variables given the available data. Parameter identification is essentially an optimization problem that aims to minimize the distance between the model-predicted

and observed (measured data) values. The problem can be formulated in the maximum likelihood approach (Walter and Pronzato, 1997) or within a Bayesian framework (Reed et al., 2016). For nonlinear problems, the optimization can result in multiple local solutions. To avoid the convergence of local solutions, global and hybrid global-local optimization methods have been developed (Walter and Pronzato, 1997; Banga, 2008). As modelers, we are interested in knowing whether the optimization problem has a unique solution. Structural identifiability analysis aims to assess the possibility of estimating a unique best value of the model parameters from available measurements. This identifiability property is of particular importance in models where the parameters have biological meaning. The evaluation of structural identifiability is only based on the mathematical structure of the model but does not depend on the actual data. This qualitative property is based on the assumption that the model is accurate (no characterization error), the measurements are noise-free exact (no measurement errors), and that the model inputs and measurement times can be chosen freely without any constraint. Accordingly, with these ideal assumptions, the property of structural identifiability does not depend on the limitations imposed by the quality and quantity of the data available for parameter identification. The rigorous mathematical framework of structural identifiability has been discussed by Walter and Pronzato (1997), while a simple introduction targeted to the animal science community was provided by Muñoz-Tamayo et al. (2018).

This paper aims to illustrate the power of identifiability analysis for developing robust predictive models. We also aimed to promote the use of identifiability analysis within the modeling construction workflow in animal science.

Defining Parameter Identifiability
Structural identifiability
Let us consider the following ODE model shown as follows:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{p}),$$

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}, \mathbf{u}, \mathbf{p}), \quad \mathbf{x}(0) = \mathbf{x}_0(\mathbf{p}) \qquad (1)$$

where $\mathbf{x}$ is the vector of state variables, $\mathbf{y}$ is the output vector (measurements), $\mathbf{u}$ is the input vector, and $\mathbf{p}$ is the parameter vector. The model structure is defined by the vector functions $\mathbf{f}$, $\mathbf{g}$ which can be linear or nonlinear. A parameter p_i is structurally identifiable if it can be uniquely recovered from information on the input and output variables. This property translates mathematically as follows:

$$\mathbf{y}(t, \hat{\mathbf{p}}) = \mathbf{y}(t, \mathbf{p}^*) \Rightarrow \hat{p}_i = p_i^* \qquad (2)$$

Structural identifiability can be local or global. The parameter p_i is structurally locally identifiable if it can be estimated in a neighborhood of its nominal value, but a finite number of possible values exist in the parameter space that holds equation (2). The parameter p_i is structurally globally identifiable if it can be uniquely estimated in the whole parameter space (Barreiro and Villaverde, 2022, preprint). If none of the previous conditions hold, the parameter p_i is non-identifiable.

It should be noted that the assessment of identifiability analysis may, in some cases, depend on the initial conditions of the state variables (Denis-Vidal et al., 2001; Saccomani et al., 2003). Indeed, in some cases, certain initial conditions may lead to the loss of identifiability. Joubert et al. (2021) proposed a method to identify some problematic initial conditions impacting parameter identifiability.

Existing methods and software tools
A variety of methods exists to test the structural identifiability of dynamic models. They include the Laplace transform (for linear models), direct test, differential algebra, Taylor series, and generating series Chis et al., 2011; Miao et al., 2011). In addition, many software tools are freely available to assess structural identifiability (locally or globally). The outcome given by software tools is qualitative; the software displays which parameters (or combination of parameters) are identifiable and which are not. Table 1 lists the availability of common and recent software tools. These software tools include DAISY (Bellu et al. 2007; Saccomani et al. 2019), ObservabilityTest (Sedoglavic 2002), IdentifiabilityAnalysis (Karlsson et al. 2012), STRIKE-GOLDD 4.0 (Villaverde et al. 2016; Díaz-Seoane et al. 2023; Hong et al. 2019), Structural Identifiability Toolbox (Ilmer et al. 2021), StrucID (Stigter and Joubert 2021), StructuralIdentifiability (Dong et al. 2022, preprint), and NonlinearObservabilityTest (Shi and Chatzis 2022). Some of these tools are implemented in commercial programming languages such as Matlab (https://fr.mathworks.com/products/matlab.html), Mathematica (https://www.wolfram.com/mathematica/), and Maple (https://www.maplesoft.com), while others are implemented in free and open-source environments such as Reduce (https://reduce-algebra.sourceforge.io/), Maxima (https://maxima.sourceforge.io/), and Julia (https://julialang.org/) (Bezanson et al., 2017). Benchmarking studies have been performed to compare identifiability software tools (Raue et al., 2014; Hong et al., 2019; Dong et al., 2022, preprint). Barreiro and Villaverde (2022, preprint) provided a comprehensive benchmarking study assessing 12 software tools for identifiability analysis. Their study discussed the strengths and weaknesses of different tools and provided software selection guidelines. For global identifiability analysis, the authors recommend using the Maple implementation of SIAN and StructuralIdentifability.

Practical identifiability
Developments in structural identifiability analysis have reached a high degree of maturity, which has led some authors to declare that determining structural identifiability is not longer an issue (Wieland et al., 2021). Since structural identifiability is a qualitative property, a quantitative assessment of the parameters' accuracy is needed to fully characterize the parameters' identifiability for a given experimental data set. This assessment is related to the notion of practical identifiability, and it should be said that structural non-identifiability implies practical non-identifiability.

Briefly, practical identifiability analysis is centered on the numerical determination of the confidence intervals of the parameter estimates. Different methods are available for the computation of parameter confidence intervals, including the Fisher information matrix (**FIM**)-based approach, Monte-Carlo simulation, Bayesian method, and profile likelihood. Practical identifiability methods were reviewed by Lam et al. (2022), and their characteristics in terms of computational

Table 1. Software tools for structural identifiability analysis

Tool	Description
DAISY	URL: https://daisy.dei.unipd.it Language: Reduce Analysis: local and global solutions Application: rational models Reference: Bellu et al. (2007); Saccomani et al. (2019)
ObservabilityTest	URL: https://github.com/sedoglavic/ObservabilityTest Language: Maple Analysis: local solution Application: rational models Reference: Sedoglavic (2002)
IdentifiabilityAnalysis	URL: https://www.fcc.chalmers.se/software/other-software/identifiabilityanalysis/ Language: Mathematica Analysis: local solution Application: rational models Reference: Karlsson et al. (2012)
STRIKE-GOLDD 4.0	URL: https://github.com/afvillaverde/strike-goldd Language: Matlab Analysis: local solution Application: rational and non-rational models Reference: Villaverde et al. (2016); Díaz-Seoane et al. (2023)
GenSSI 2.0	URL: https://github.com/genssi-developer/GenSSI Language: Matlab Analysis: local and global solutions Application: rational and non-rational models Reference: Chiş et al. (2011); Ligon et al. (2018)
COMBOS	URL: http://biocyb1.cs.ucla.edu/combos/ Language: Maxima, web application Analysis: local and global solutions Application: rational models Reference: Meshkat et al. (2014)
SIAN	URL: https://github.com/alexeyovchinnikov/SIAN-Julia https://github.com/pogudingleb/SIAN Language: Maple and Julia Analysis: local and global solutions Application: rational models Reference: Hong et al. (2019)
Structural Identifiability Toolbox	URL: https://maple.cloud/app/6509768948056064 Language: Maple, web application Analysis: local and global solutions Application: rational models Reference: Ilmer et al. (2021)
StrucID	URL: available upon request from the authors Language: Matlab Analysis: local solution Application: rational and non-rational models Reference: Stigter and Joubert (2021)
StructuralIdentifiability	URL: https://github.com/SciML/StructuralIdentifiability.jl Language: Julia Analysis: local and global solutions Application: rational models Reference: Dong et al. (2022, Preprint)
NonlinearObservabilityTest	URL: https://eng.ox.ac.uk/non-lineardynamics/resources/ Language: Matlab Analysis: local solution Application: rational models Reference: Shi and Chatzis (2022)

cost and statistical interpretability were discussed by Villaverde et al. (2022). Software tools for parameter identification allow for practical identifiability analysis based on the FIM (Muñoz-Tamayo et al., 2009; Balsa-Canto et al., 2016), sensitivity analysis (Soetaert and Petzoldt, 2010), or profile likelihood approach (Raue et al., 2015).

It is important to emphasize that the conceptual development of practical identifiability analysis is less mature than structural identifiability (Wieland et al., 2021). The logic sequence between the two identifiability notions of data collection explains that structural identifiability is also called a priori identifiability, while practical identifiability is termed

a posteriori identifiability. The joint integration of structural and practical identifiability analyses offers a powerful armory to tackle the parameter identification of models of biological systems (Miao et al., 2011; Saccomani and Thomaseth, 2018).

In addition to structural and practical identifiability analyses, sensitivity analysis can provide helpful information on parameter identifiability. A sensitivity analysis study allows for assessing how the model outputs are affected by different sources of variation, including the model parameters (Saltelli et al., 2000). That is, how the change of a parameter impacts the behavior of the model output. Sensitivity analysis is central to identifying the phenomena that play a significant role in system behavior and ranking the model parameters regarding their influence on the model outputs. Various model developments in animal science include sensitivity analysis to evaluate the effect of variation of parameters and input variables on the model behavior (Doeschl-Wilson et al., 2009; Tedeschi and Fox, 2009; Puillet et al., 2016; Dougherty et al., 2017; van Lingen et al., 2019).

The FIM-based approach for practical identifiability is based on the calculation of the sensitivity of the model to its parameters. Indeed, if the model outputs are highly sensitive to a small perturbation of a given parameter, this parameter is likely to be identifiable (Miao et al., 2011). The information provided by sensitivity analysis is useful, for example, to discard parameters with little influence on the model outputs and to reduce the number of parameters estimated with the model calibration.

Parameter Identifiability in Animal Science

In the animal nutrition field, the concept of structural identifiability was introduced by Boston et al. (2007) and further detailed by Tedeschi and Boston (2010), focusing on linear models. Later on, Muñoz-Tamayo et al. (2018) expanded on the mathematical elements and notions associated with the

structural identifiability analysis for nonlinear ODE models and discussed the relevance of identifiability analysis in the modeling construction. The paper illustrates relevant examples that demonstrate the consequences of ignoring identifiability testing when the identifiability question is relevant. It also shows the implications that can have the use of identifiability on the costs of the experimental design. To emphasize the relevance of identifiability analysis, let us consider a compartmental model (equations (3) and (4)) that is prevalent in animal science, and their applications include, for example, the study of rumen digesta passage (Poppi et al., 2001) and physiological regulations (Puillet et al., 2008).

$$\frac{dx_1}{dt} = -p_1 \cdot x_1 \tag{3}$$

$$\frac{dx_2}{dt} = p_1 \cdot x_1 - p_2 \cdot x_2 \tag{4}$$

This compartmental model has two parameters p_1, p_2 and two state variables x_1, x_2. We consider that only x_2 can be measured, and the initial condition $x_1 (0)$ is unknown. Under this condition, the parameters are locally identifiable. The results using DAISY are displayed in Figure 1. For the sake of clarity, it should be said that the algorithm implemented in DAISY uses random numerical values for the parameters to resolve the system of equations of the identifiability test. These random numerical values are those displayed in Figure 1. They are not the actual values of the parameters. As previously mentioned, the numerical values of the parameters are obtained from a parameter identification routine that minimizes the distance between the model-predicted and observed values. We observed that DAISY gives two solutions for the identifiability where the values of the parameters are switched between the two solutions. That means that if the true value of the parameters are p_1^*, p_2^*, the correct solution


```
§  0.13+0.27 secs            reduce
File  Edit  Font  Break  Load Package  Switch


MODEL EQUATION(S)$

c_  := {df(x1,t)= - p1*x1,df(x2,t)=p1*x1 - p2*x2,y1=x2}$

         ⋮
         ⋮

RANDOMLY CHOSEN NUMERICAL PARAMETER(S) VECTOR$

b2_  := {p1=33,p2=38}$

EXHAUSTIVE SUMMARY $

flist_  := {p1*p2 - 1254,p1 + p2 - 71}$

MODEL PARAMETER SOLUTION(S)$

 G_:=GROESOLVE(FLIST_,B1_) $

g_  := {{p1=38,p2=33},{p1=33,p2=38}}$

MODEL LOCALLY IDENTIFIABLE$
```

Figure 1. Scheme of a compartmental model and the output of identifiability testing with the DAISY software (Bellu et al. 2007). If only x_2 is measured and the initial condion $x_1(0)$ is unknown the parameters are locally identifiable. Two solutions exist where the values of the parameters are switched between the solutions.

is $p_1 = p_1^*; p_2 = p_2^*$, while the second (wrong) solution is $p_1 = p_2^*; p_2 = p_1^*$. This outcome results from the fact that, to solve the parameter estimation problem, the initial condition $x_1(0)$ should also be considered as a parameter to be estimated because its value is unknown. The solutions found by DAISY are associated with two different $x_1(0)$. Let's assume that the modeler has a perfect data set. The optimization algorithm resolves the wrong parameter values when performing the calibration procedure. The modeler, unaware of the identifiability analysis, could conclude that the parameter estimation is satisfactory because the model fitting is perfect. What are the consequences of this situation? Firstly, the parameters p_1, p_2 might have biological meaning, so conclusions based on the specific values of the parameter estimates will be wrong. Note, for example, that p_1 and p_2 can have a different order of magnitude, which is a typical case of systems where internal processes occur at different time scales (i.e., slow vs. fast processes). The second consequence is that with the wrong parameter estimates, the model can represent very well the dynamics of x_2. However, the model will fail to provide accurate predictions of x_1. Many model developments aim at providing predictions of unobserved variables; therefore, it is essential to estimate accurate parameters. An identifiability analysis can help to fulfill this goal. In the example, to obtain a unique set of parameters, we only need to know the value of the initial condition $x_1(0)$. Note that this example is not the worst-case scenario that the modeler can find because the parameters are at least locally identifiable. The most critical situation is when the parameters are non-identifiable. Muñoz-Tamayo et al. (2018) illustrate an example where the wrong choice of the initial condition can render the parameters non-identifiable. The identifiability of the parameters is guaranteed by simply changing the initial conditions. This capability that the modeler has to set the conditions of an experimental protocol to ensure parameter identifiability before data collection refers to the room of maneuver mentioned above.

With the previous example, we briefly showed the relevance of identifiability analysis. Despite our motivation to promote structural identifiability analysis in the modeling arena in animal science, this approach is still seldom applied to the study of dynamic models. Therefore, in our paper, we focus on dynamic models described by ODE. However, identifiability analysis also applies to statistical models. Examples of identifiability analysis of statistical models used in animal genetics have also been addressed (Cantet and Cappa, 2008; Shariati et al., 2009).

Few dynamic modeling studies integrate structural identifiability analysis. These studies include a model for the transmission of mastitis in dairy cows (White et al., 2002), a Gompertz-based model to describe the body weight dynamics of piglets at weaning (Revilla et al., 2019), a model to quantify the response of feed intake of pigs facing a perturbation (Nguyen-Ba et al., 2020), a model to characterize body condition score variations in sheep (Macé et al., 2020, Preprint), a model to describe the methanogenesis by rumen archaea (Muñoz-Tamayo et al., 2019a), and a model to describe in vivo methane production from cattle (Muñoz-Tamayo et al., 2019b).

Although structural identifiability has been rarely applied in the animal science field, considerations of the practical and numerical issues of the model calibration are obliged aspects that modelers face to find an adequate strategy that facilitates the numerical estimation of the model parameters. For example, sensitivity analysis was applied to two mathematical models developed to describe the susceptibility of porcine alveolar macrophages to an RNA virus (Doeschl-Wilson et al., 2016). This approach allows for a reduction in the number of parameters identified by fixing the values of a subset of the parameters. Similarly, sensitivity analysis was used to perform a stepwise fitting procedure to estimate the parameters of a model of the bovine estrous cycle (Boer et al., 2017). A two-step parameter identification strategy to limit practical identifiability issues was implemented to identify the parameters of a lactation model that account for perturbations (Ben Abdelkrim et al., 2021). In a modeling development describing the interaction between the growth rate of the developing embryo and the uterine environment in cows, identifiability analysis was performed to guarantee the unicity of the parameters (Shorten et al., 2018). However, no details were provided about how the identifiability analysis was done. We might infer that the authors refer to practical identifiability analysis extracted from the standard error calculation by the Markov chain Monte-Carlo method. The profile likelihood approach (Murphy and Van Der Vaart, 2000; Raue et al., 2009) was applied to assess the practical identifiability of a model developed to describe the effect of diet composition on sheep weight (Vargas-Villamil et al., 2020).

Our objective in this paper follows up on our previous attempts to promote identifiability analysis in the modeling exercise in our field. We will illustrate the power of identifiability analysis using models at different levels of complexity. Following open science practices (Muñoz-Tamayo et al., 2022), the source codes with the implementation of identifiability analysis for all the examples are freely available (Muñoz-Tamayo, 2023).

The Power of Model Simplification

We consider here the process of protein hydrolysis in the context of cheese ripening. In this hydrolytic process, lactic acid bacteria break down milk proteins (e.g., β-casein) into various peptides, which are further metabolized. The following model describes the hydrolysis of the intact β-casein by the P_I-type protease of *Lactococcus lactis* in a batch system (Muñoz-Tamayo et al., 2011).

$$\frac{dx}{dt} = -k \cdot E \cdot \frac{x}{K_m \cdot (1 + (I/k_I)) + x} \tag{5}$$

where x is the concentration of β-casein (mol/L), E (mol/L) is the concentration of the protease, I (mol/L) is the concentration of the inhibitor compound(s), k is the catalytic rate constant (mol x/(mol E min)), K_m (mol/L) is the affinity substrate constant, and k_I is the inhibition constant (mol/L). The model in equation (5) can be categorized as a mechanistic model. It is derived from a mass balance, and its parameters are biologically meaningful (interpretable).

Let us consider that only the concentration β-casein (mol/L) was measured at different sampling times. The concentration of the enzyme E is known and constant ($E = 1$). We will then need information about the inhibitor I to set up the parameter identification problem. During the hydrolysis, there is a competition between the intact protein and the released peptides for the active sites of the protease. Indeed, the kinetic function in Eq. (5) is called a competition inhibition kinetic rate. We can then consider that the inhibitor is the sum of all peptides

released. Accordingly, $I = x_0 - x$, where x_0 is the initial concentration of β-casein. The initial concentration x_0 is known ($x_0 = 10$).

Identifiability analysis was done with DAISY, Nonlinear-ObservabilityTest, COMBOS, GenSSI 2.0, STRIKE-GOLDD 4.0, StructuralIdentifiability, and SIAN (implemented in Julia). The results led to the conclusion that the parameters k, K_m, k_I are non-identifiable. Figure 2 shows the outcome of the analysis using StructuralIdentifiability. This result should not discourage us. We are actually in a very typical situation of over-parameterization (too many parameters). If we have information on prior values of any of the three parameters, we can set the parameter as known and let the other remaining parameters be estimated. Fixing the value of one parameter as known will render the other parameters globally identifiable. The previous solution directly reduces the number of parameters using prior knowledge of the parameter values. However, what happens if we do not have any prior information on any of the three parameters or if we set an incorrect prior value? A solution still exists; it is to re-parameterize the model. The model of this example is simple, and we can engage the parameterization by hand. By manipulation Eq. (5), we obtain the following reduced model, shown in Eqs (4)-(6)
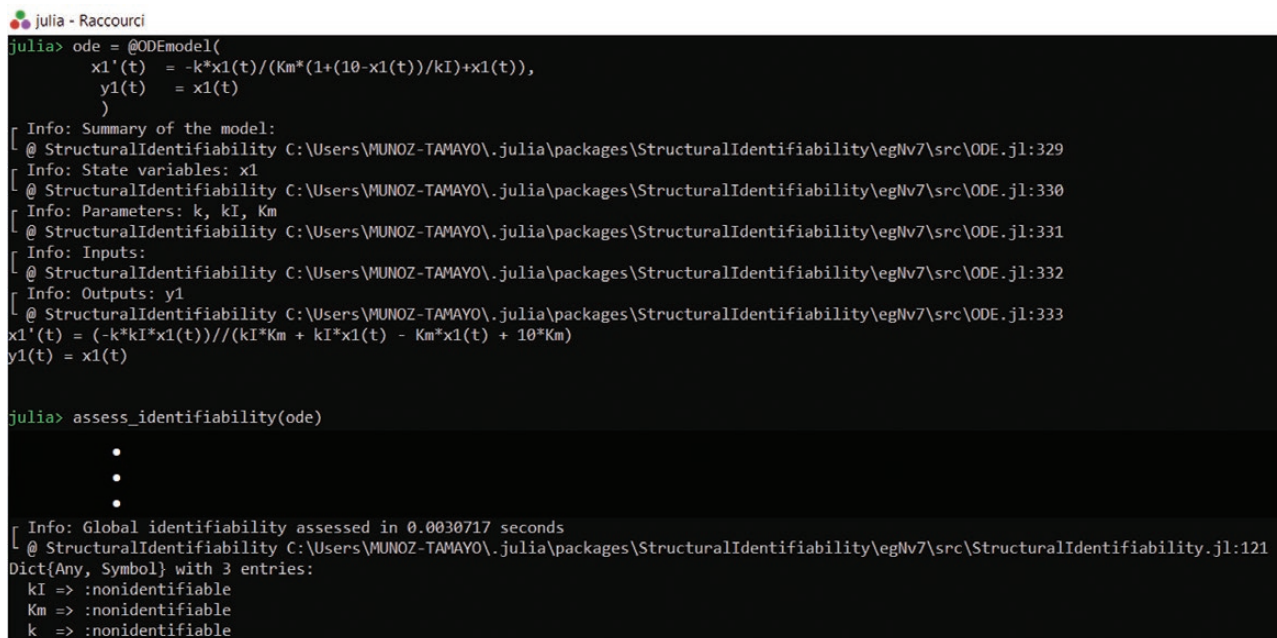
$$\frac{dx}{dt} = -b_1 \cdot E \cdot \frac{x}{b_2 - x} \qquad (6)$$

with

$$b_1 = \frac{k \cdot k_I}{K_m - k_I} \qquad (7)$$

$$b_2 = \frac{K_m \cdot (k_I + x_0)}{K_m - k_I} \qquad (8)$$

The parameters b_1, b_2 are globally identifiable. The reparameterization helps here to improve the identifiability properties of the model. On the other hand, we lose parameter interpretability (Lema-Perez et al., 2019). The reparameterization task within the model-building process is indeed an exercise of trade-offs. In this simple example, the reparameterization can be done by inspection. However, for more complex models, the reparameterization can be challenging to reach by simple inspection. The Matlab application StrucID (Stigter and Joubert, 2021) allows the detection of the lack of identifiability in ODE models. The analysis provides information on correlations between potential non-identifiable parameters. This information can be further used within the procedure developed by Joubert et al. (2020) to obtain suitable reparameterizations to improve the identifiability of the model. The reparameterization process is, however, complicated and requires expert knowledge of mathematics and computer programming. In this regard, the COMBOS, STRIKE-GOLDD 4.0 tools, and the web application developed by Ilmer et al. (2021) provide helpful functionality since they allow the computation of identifiable combinations of parameters that are individually non-identifiable. These combinations can indeed inform reparameterizations for model simplification and to guarantee structural identifiability. Moreover, in some cases, the resulting identifiable combinations can have biological meaning for the system under study (Ilmer et al., 2021). The automatic reparameterization in STRIKE-GOLDD 4.0 is performed by the implementation of the AutoRepar procedure (Massonis et al., 2021). The web tool web application developed by Ilmer et al. (2021) uses the SIAN algorithm (Hong et al., 2019) for identifiability testing and the algorithm developed by Ovchinnikov et al. (2021) for computing identifiable combinations. GenSSI 2.0 has an implementation of parameter transformation to facilitate the removal of non-identifiable parameters.


```
julia - Raccourci
julia> ode = @ODEmodel(
        x1'(t)   = -k*x1(t)/(Km*(1+(10-x1(t))/kI)+x1(t)),
        y1(t)    = x1(t)
        )
[ Info: Summary of the model:
  @ StructuralIdentifiability C:\Users\MUNOZ-TAMAYO\.julia\packages\StructuralIdentifiability\egNv7\src\ODE.jl:329
[ Info: State variables: x1
  @ StructuralIdentifiability C:\Users\MUNOZ-TAMAYO\.julia\packages\StructuralIdentifiability\egNv7\src\ODE.jl:330
[ Info: Parameters: k, kI, Km
  @ StructuralIdentifiability C:\Users\MUNOZ-TAMAYO\.julia\packages\StructuralIdentifiability\egNv7\src\ODE.jl:331
[ Info: Inputs:
  @ StructuralIdentifiability C:\Users\MUNOZ-TAMAYO\.julia\packages\StructuralIdentifiability\egNv7\src\ODE.jl:332
[ Info: Outputs: y1
  @ StructuralIdentifiability C:\Users\MUNOZ-TAMAYO\.julia\packages\StructuralIdentifiability\egNv7\src\ODE.jl:333
x1'(t) = (-k*kI*x1(t))//(kI*Km + kI*x1(t) - Km*x1(t) + 10*Km)
y1(t) = x1(t)


julia> assess_identifiability(ode)
        •
        •
        •
[ Info: Global identifiability assessed in 0.0030717 seconds
  @ StructuralIdentifiability C:\Users\MUNOZ-TAMAYO\.julia\packages\StructuralIdentifiability\egNv7\src\StructuralIdentifiability.jl:121
Dict{Any, Symbol} with 3 entries:
  kI  => :nonidentifiable
  Km  => :nonidentifiable
  k   => :nonidentifiable
```

Figure 2. Output of identifiability testing of the model of β-casein hydrolysis with the StructuralIdentifiability software (Dong et al. 2022, preprint). The result of identifiability analysis is qualitative and gives information on which parameters are identifiable and which are not. In the example, all the parameters are non-identifiable.

The Power of Selecting What to Measure

The following example uses a mathematical model describing methane production (CH_4) by rumen methanogenic archaea in an in vitro batch system (Muñoz-Tamayo et al., 2019a). Equations (9) to (13) show the model equations.

$$\frac{dx_{H_2}}{dt} = \mu_{max} \cdot \exp\left(-\frac{K_s \cdot V_g}{n_{g,H_2}}\right) \cdot x_{H_2} - k_d \cdot x_{H_2} \tag{9}$$

$$\frac{ds_{CO_2}}{dt} = -\frac{-Y_{CO_2} \cdot \mu_{max}}{Y} \cdot \exp\left(-\frac{K_s \cdot V_g}{n_{g,H_2}}\right) \cdot \\ x_{H_2} - k_L a \cdot \left(s_{CO_2} - K_{H,CO_2} \cdot R \cdot T \cdot n_{g,CO_2}/V_g\right) \tag{10}$$

$$\frac{dn_{g,H_2}}{dt} = -\frac{\mu_{max}}{Y} \cdot \exp\left(-\frac{K_s \cdot V_g}{n_{g,H_2}}\right) \cdot V_L \cdot x_{H_2} \tag{11}$$

$$\frac{dn_{g,CO_2}}{dt} = V_L \cdot k_L a \cdot \left(s_{CO_2} - K_{H,CO_2} \cdot R \cdot T \cdot n_{g,CO_2}/V_g\right) \tag{12}$$

$$\frac{dn_{g,CH_4}}{dt} = \frac{Y_{CH_4} \cdot \mu_{max}}{Y} \cdot \exp\left(-\frac{K_s \cdot V_g}{n_{g,H_2}}\right) \cdot V_L \cdot x_{H_2} \tag{13}$$

where s_{CO_2} is the concentration (mol/L) of carbon dioxide in the liquid phase, and x_{H_2} is the biomass concentration (mol/L) of hydrogenotrophic methanogens. The number of moles in the gas phase is represented by the variables $n_{g,H_2}, n_{g,CO_2}, n_{g,CH_4}$. The constants V_g and V_L are the volume of the gas and liquid phase, respectively. Liquid-gas transfer for carbon dioxide is determined by the mass transfer coefficient $k_L a$ (h^{-1}) and Henry's law coefficient K_{H,CO_2}(M/bar). The constant R (bar $(M \cdot K)^{-1}$) is the ideal gas law constant, and T is the temperature (K). The parameter k_d (h^{-1}) is the death cell rate constant. The parameters Y, Y_{CO_2}, Y_{CH_4} are the yield factors (mol/mol H_2) of microbial biomass, CO_2, and CH_4 that account for the stoichiometry of the reactions. The model uses the microbial growth function proposed by Desmond-Le Quemener and Bouchez (2014), with hydrogen as the limiting substrate.

$$\mu = \mu_{max} \cdot \exp\left(-\frac{K_s \cdot V_g}{n_{g,H_2}}\right) \tag{14}$$

where μ is the growth rate (h^{-1}), μ_{max} (h^{-1}) is the maximum specific growth rate constant and K_s(mol/L) the affinity constant. An implementation of the model in the Open Source software Scilab is available at https://doi.org/10.5281/zenodo.3271611.

Let us assume that only the concentration of methanogens x_{H_2} and the moles of hydrogen n_{g,H_2} are measured. We are interested in assessing the identifiability of the biological parameters $\mu_{max}, K_s, k_d, Y, Y_{CO_2}, Y_{CH_4}$. All initial concentrations are known. The remaining (physical-related) parameters are known.

The previous model is non-rational since it includes an exponential function. Identifiability analysis was done with GenSSI 2.0 and STRIKE-GOLDD 4.0, which are a few tools that can analyze non-rational models. Under the observation conditions, the parameters μ_{max}, K_s, k_d, Y are globally identifiable, while Y_{CO_2}, Y_{CH_4} are non-identifiable. The result of the non-identifiable parameters is not surprising. We can check by inspection of the model equations that it will be impossible to estimate the parameters Y_{CO_2}, Y_{CH_4} without measuring, respectively, CO2 and CH4. Indeed, we need the information on these quantities to estimate the relation of moles consumed or produced by mole of H2 utilized. Thus, the complete set of parameters is globally identifiable if $x_{H_2}, n_{g,H_2}, n_{g,CO_2}, n_{g,CHO_4}$ are measured.

We can continue our analysis to illustrate the importance of integrating biological knowledge into the model. Methanogenesis is a process involving methane and microbial biomass production. We can represent the process in two reactions:

$$R_1 : 4H_2 + CO_2 \rightarrow CH_4 + 2H_2O$$

$$R_2 : 10H_2 + 5CO_2 + NH_3 \rightarrow C_5H_7O_2N + 8H_2O$$
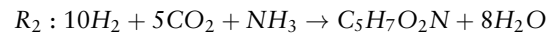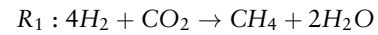
where $C_5H_7O_2N$ is the chemical formula for microbial biomass. Knowing the stoichiometry of the reactions enables us to reduce the number of yield parameters. The yield factor Y is the number of moles of microbial biomass produced per mole of H_2 via reaction R_2. We can then express the fraction (f) of H_2 utilized in reaction R_1 for methane production as a function of Y:

$$f = 1 - 10 \cdot Y \tag{15}$$

The number 10 in Equation (15) is the stoichiometry coefficient of H_2 in R_2. The yield factors of CO_2 and CH_4 can now be expressed as functions of f:

$$Y_{CO_2} = \left(\frac{1}{4}\right) \cdot f + \left(\frac{5}{10}\right) \cdot (1-f) \tag{16}$$

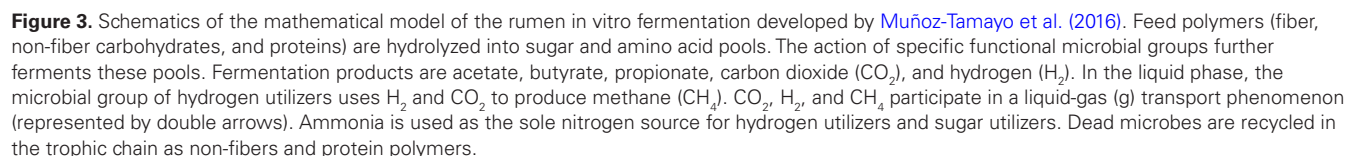$$Y_{CH_4} = \left(\frac{1}{4}\right) \cdot f \tag{17}$$

This means that the number of parameters is now reduced to four parameters instead of 6. All the parameters are identifiable under the scenario where x_{H_2} and n_{g,H_2} are measured

The Power of Analyzing Complex Models

With the recent progress in computational methods, structural identifiability testing (at least locally) can be applied to complex nonlinear models. For example, previous studies (Ligon et al., 2018; Barreiro and Villaverde, 2022, preprint) showed that SIAN, StructuralIdentifiability, IdentifiabilityAnalysis, and GenSSI 2.0 were able to test the identifiability of models with more than 20 states variables and 20 parameters. IdentifiabilityAnalysis was used to assess the local structural identifiability of a model with about 100 states and 100 parameters (Karlsson et al., 2012). In the following example, we consider a mathematical model that represents the rumen fermentation under in vitro conditions (Muñoz-Tamayo et al., 2016). Figure 3 shows the schematics of the model. The model has 18 state variables and 30 parameters. A Matlab implementation of the model is freely available at https://doi.org/10.5281/zenodo.4047640. An implementation is also available in the R-package microPop

(Kettle et al., 2018). An implementation in Scilab is also available at https://doi.org/10.5281/zenodo.4090332 for an extended model that accounts for the effect of the macroalgae *Asparagopsis taxiformis* on rumen fermentation and methane production (Muñoz-Tamayo et al., 2021). The original model includes algebraic equations to compute the pH dynamically. For our identifiability exercise, we set the pH to a constant value of 6.6.

We will consider the following 14 parameters for identifiability analysis: $k_{hyd,ndf}$ (hydrolysis rate constant of cell wall carbohydrates), $k_{hyd,nsc}$ (hydrolysis rate constant of non-structural carbohydrates), $k_{hyd,pro}$ (hydrolysis rate constant of proteins), $k_{m,su}$, (maximum specific utilization rate constant of amino sugars), $K_{s,su}$ (substrate affinity constant of sugars), Y_{su} (microbial yield factor of sugars utilizers), $k_{m,aa}$ (maximum specific utilization rate constant of amino acids), $K_{s,aa}$,(substrate affinity constant of amino acids), Y_{aa} (microbial yield factor of amino acids utilizers), k_{m,H_2}(maximum specific utilization rate constant of hydrogen), K_{s,H_2} (substrate affinity constant of hydrogen utilization, Y_{H_2} (microbial yield factor of hydrogen utilizers), and λ_1, λ_2 (flux distribution parameters).

The initial conditions were set to be known. Identifiability testing was done with StructuralIdentifiability, GenSSI 2.0, and STRIKE-GOLDD 4.0. We run the tests on a laptop with Windows 64 Gb RAM, Intel Core i9-10885H (8 cores, 2.4 GHz). We first considered that 12 state variables were observed. The remaining unobserved state variables were the concentrations of CO_2, H_2, CH_4 in the liquid phase and the

concentration of the three microbial functional groups. The runtimes for the local structural identifiability analysis were 2.5 s for STRIKE-GOLDD 4.0, and 4.2 s for StructuralIdentifiability. The runtimes for global structural analysis were 2.6 min for StructuralIdentifiability and 25 min for GenSSI 2.0. Under the tested condition, the parameters are globally identifiable.

In the paper by Muñoz-Tamayo et al. (2016), the parameter estimation was defined for a subset of 10 parameters: $k_{hyd,nsc}$, $k_{hyd,pro}$, $k_{m,su}$, Y_{su}, $k_{m,aa}$, Y_{aa}, k_{m,H_2}, Y_{H_2}, λ_1, λ_2 . The remaining model parameters were fixed as known. This strategy was meant to facilitate the model calibration routine. The observed variables were the concentrations of acetate, butyrate, propionate, ammonia, and the moles of H_2, CO_2, and CH_4 in the gas phase. When the model was built, no structural identifiability analysis was done. In the present exercise, we used StructuralIdentifiability and GenSSI 2.0 for identifiability testing. Both tools could not test the global identifiability of the parameters under the observation conditions. The analysis with both tools indicated that the model parameters are locally identifiable. The runtime with StructuralIdentifiability was 3.0 s. The runtime with GenSSI 2.0 was 34 min, but this time included the test on global identifiability, which was unsuccessful. STRIKE-GOLDD 4.0 exhibited an error in the process and could not assess the identifiability. We further evaluated the local structural identifiability of the 14 model parameters under the most restricted observation condition (that is, one single measurement). The result with StructuralIdentifiability informed that measuring any volatile

Figure 3. Schematics of the mathematical model of the rumen in vitro fermentation developed by Muñoz-Tamayo et al. (2016). Feed polymers (fiber, non-fiber carbohydrates, and proteins) are hydrolyzed into sugar and amino acid pools. The action of specific functional microbial groups further ferments these pools. Fermentation products are acetate, butyrate, propionate, carbon dioxide (CO_2), and hydrogen (H_2). In the liquid phase, the microbial group of hydrogen utilizers uses H_2 and CO_2 to produce methane (CH_4). CO_2, H_2, and CH_4 participate in a liquid-gas (g) transport phenomenon (represented by double arrows). Ammonia is used as the sole nitrogen source for hydrogen utilizers and sugar utilizers. Dead microbes are recycled in the trophic chain as non-fibers and protein polymers.

fatty acid (acetate, butyrate, and propionate) yielded local identifiability of the parameters. This result is encouraging for the rumen modelers, although we might recognize that this outcome might appear surprising. The reason for the identifiability of the model parameters is associated with the nonlinear structure of the model. Although model complexity is not a reliable indicator to compare the identifiability properties between models (Roper et al., 2010), nonlinear complex models are likely more identifiable than linear models (Walter and Pronzato, 1996).

The identifiability results of the rumen fermentation model developed by Muñoz-Tamayo et al. (2016) are similar to those found by Nimmegeers et al. (2017) when studying the identifiability of the Anaerobic Digestion Model No.1 (ADM1). The ADM1 (Batstone et al., 2002) is a model developed to represent the digestion in reactors for wastewater treatment. The rumen fermentation model has a similar structure to that of ADM1, which explains the structural identifiability property shared between the models. For both models, a minimal set of measurements can guarantee local structural identifiability. Nimmegeers et al. (2017) explained that structural identifiability results from many interconnections between the model's state variables. Such interconnection, however, also applies to the parameter set, leading to practical identifiability issues. Optimal experiment design can help to remediate practical identifiability issues, as discussed in the next section.

This example shows that existing identifiability analysis tools allow for handling complex models. Although for some models, it may not be possible to perform parameter identifiability analysis. What to do in these cases? We discuss some solutions here below.

What can we do when identifiability testing is not possible?

Although developing advanced tools for structural identifiability testing might occur for some models with high complexity and limited observation conditions, current software tools cannot solve the identifiability problem. As mentioned above, from the tools shown in Table 1, only GenSSI 2.0 and STRIKE-GOLDD 4.0 can analyze non-rational models using symbolic computation. This can be a limitation issue in animal science, where non-rational functions (e.g.,

exponential functions) are common. Although, in some cases, transformations are possible to render the non-rational model in polynomial or rational form (see Muñoz-Tamayo et al., 2018 for an example of a transformation).

When identifiability tools fail to assess identifiability, the numerical approach implemented in the StrucID app can help identify correlations between potential non-identifiable parameters quickly. This valuable information can be used further to refine models and reduce the set of parameters to be checked for identifiability. The method implemented in StrucID app is not restricted to rational models. The methods developed within the practical identifiability framework (Lam et al., 2022; Villaverde et al., 2022) also provide valuable resources to assess a posteriori identifiability when structural identifiability testing is out of reach.

Finally, it is important to point out that automatic methods for identifiability testing are not free of error. In some cases, identifiability tools can yield incorrect results (Dong et al., 2022, Preprint). It is thus advisable to use different tools simultaneously to assess the correctness of the results (Joubert et al., 2021; Barreiro and Villaverde, 2022, preprint).

The Power of Designing Optimal Experiments

Parameter identification is often addressed like a downstream process after collecting data. By following this approach, the modeler has minimal room for maneuvering to improve the model's accuracy. The accuracy measures how closely the model predictions are to the actual values (Tedeschi, 2006). By incorporating identifiability analysis, we can follow an upstream approach to increase the room for maneuver of the modeler in the modeling construction process. The previous examples illustrate that the first benefit of identifiability analysis is providing valuable information about what to measure to render the model parameters identifiable. This part is done within the framework of structural identifiability. To complete the picture, we will need to know under which experimental conditions the measurements should be done on the real system to guarantee accurate parameter estimates. This part is addressed by practical identifiability analysis. To illustrate the usefulness of practical identifiability for optimal experiment design (**OED**) for parameter estimation, let us consider the
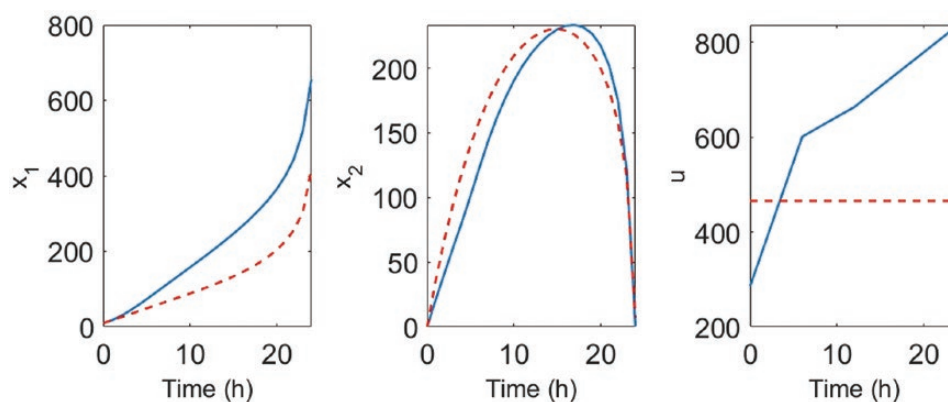

Figure 4. Responses of the variables of a mathematical model of microbial growth under two inputs of substrate concentration (u) were obtained to maximize the accuracy of the parameter estimates within an optimal experiment design. Dashed lines are the responses considering constant substrate input u_c and solid lines are the responses for a dynamic substrate input u_d.

Table 2. Comparison of the two optimal substrate inputs on the accuracy of the estimates of a model of microbial growth

	Constant input substrate u_c	Dynamic input substrate u_d
Determinant of the FIM	1.27×10^{10}	8.40×10^{12}
SD k	0.0247	0.0076
SD k_I	0.0282	0.0202

FIM, Fisher information matrix.

following model that represents the utilization of a substrate x_2 by a microbe x_1 in a continuous system as follows:

$$\frac{dx_1}{dt} = \frac{x_1 \cdot x_2}{x_2 + k + x_2^2/k_I} - D \cdot x_1 \tag{18}$$

$$\frac{dx_2}{dt} = -\frac{x_1 \cdot x_2}{x_2 + k + x_2^2/k_I} + D \cdot (u - x_2) \tag{19}$$

where D is the known dilution rate of the system, and u is the input substrate concentration. We would like to determine
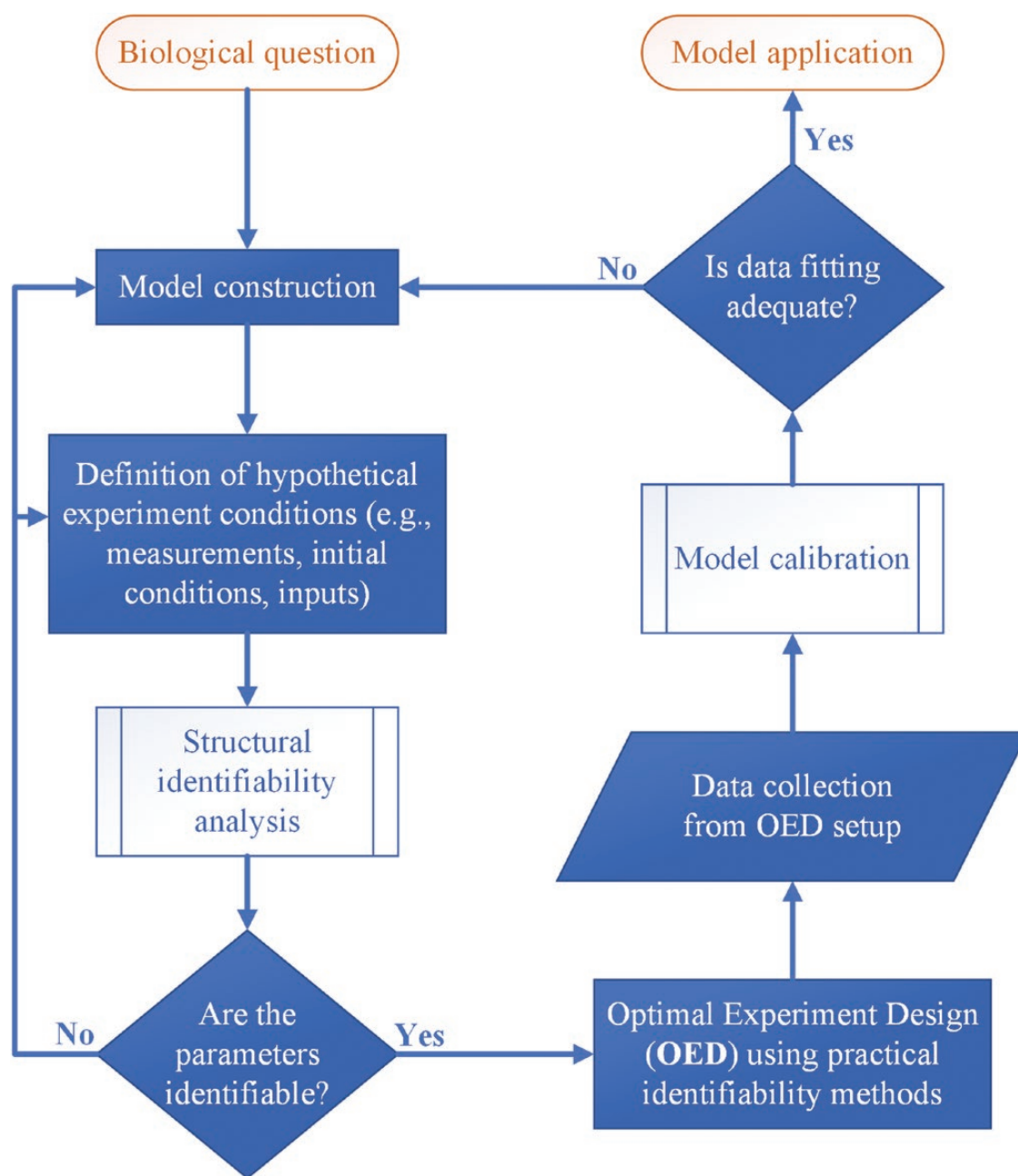
Figure 5. Workflow of the upstream approach proposed in this work. Structural and practical identifiability analysis and practical identifiability analysis are used before data collection to inform optimal experimental conditions. In the iterative process, experimental data and the mathematical model structure are modified to ensure that the model predictions are reliable.

the shape of u to perform an experiment that allows estimating accurately the parameters k, k_I from available measurements of x_1, x_2. Under these observation conditions, the parameters are structurally globally identifiable. The accuracy of the estimates translates into small confidence intervals. Our objective function can be set up as an optimization problem where we want to find u such that the volume of the confidence intervals is minimized. One approach to address the OED problem is maximizing the determinant of the FIM. Muñoz-Tamayo et al. (2018) discussed the details of the calculation of the FIM and its use in the OED. This procedure requires defining nominal values for the parameters. For our example, we set $k = 2, k_I = 50$. We solve the OED problem for two cases. In the first one, we considered the input substrate concentration to be constant over time (u_c). In the second case, we considered the input substrate concentration to vary on time (u_d). For that we parameterized u_d as a piecewise linear function.

The OED problem requires an intermediate level in computer programming skills. We used the IDEAS toolbox (Muñoz-Tamayo et al., 2009) to generate the functions for the OED problem. IDEAS is freely available at http://genome.jouy.inra.fr/logiciels/IDEAS. The files for the OED example are available at https://github.com/rafaelmunoztamayo/identifiability_examples. The toolbox Amigo 2 (Balsa-Canto et al., 2016) has the functionality of addressing OED problems for parameter estimation.

Figure 4 shows the obtained optimal inputs and the responses of the model variables. It is challenging to draw conclusions of this figure. However, when we look at the standard deviations (**SD**) of the parameter estimates obtained from the two cases in Table 2, we can clearly see the difference. The determinant of the FIM for u_d is 660 times higher than for u_c, which translates into smaller confidence intervals. The standard deviations obtained with u_d are 31% and 72% those obtained with u_c for k and k_I, respectively. Dynamic input induces better stimuli to system behavior and thus results in data with higher informative content than those obtained with constant input. This example shows the usefulness of OED in producing informative data for parameter estimation, and this capability can be used to avoid useless experiments. It should be said, however, that using the FIM for the calculation of confidence intervals is valid under linearity and asymptotic conditions (Walter and Pronzato, 1997). Approaches like the profile likelihood allow for overcoming the shortcomings of the FIM-based approach (Wieland et al., 2021).

Summary

In this paper, we showed that using an upstream approach that incorporates identifiability analysis before data collection in the workflow of model construction provides substantial benefits to obtaining reliable models. This workflow is displayed in Figure 5. Structural and practical identifiability analyses inform the conditions required to guarantee a unique and accurate estimation of the parameters. In case of lack of identifiability, identifiability analysis provides valuable information on possible actions to cure the non-identifiability (when possible). This information includes model reduction, reparameterization, and specifications on optimal measurements.

Existing freely available software tools enable the application of structural identifiability analysis without needing to be

an expert in mathematics and computer programming. Recent software tools for structural identifiability analysis allow for handling complex models, but identifiability testing might be out of reach in some cases. In this case, numerical approaches within the practical identifiability framework can address the identifiability question. We believe that this paper will motivate the modeling community in animal science to integrate identifiability analysis in their model developments. Such integration can be easily done following a practitioner approach taking advantage of the variety of available software tools dedicated to identifiability testing. However, we must stress that the practitioner approach advocated in this paper is only possible thanks to the open science practices adopted by the parameter identifiability community in making their software toolboxes freely available. The parameter identifiability topic is an excellent example of how adopting open science practices can contribute to scientific progress. We, animal scientists, should learn from such efforts to make Open Science the new normal in our field (Muñoz-Tamayo et al., 2022). By sharing data, code scripts, and software tools and making our research freely accessible, we substantially strengthen the scientific progress of the animal science domain.

Conflict of Interest Statement

The authors declare no conflicts of interest.

Literature Cited

Ben Abdelkrim, A. B., L. Puillet, P. Gomes, and O. Martin. 2021. Lactation curve model with explicit representation of perturbations as a phenotyping tool for dairy livestock precision farming. Animal. 15:100074. doi:10.1016/J.ANIMAL.2020.100074

Balsa-Canto, E., D. Henriques, A. Gábor, and J. R. Banga. 2016. AMIGO2, a toolbox for dynamic modeling, optimization and control in systems biology. Bioinformatics. 32:3357–3359. doi:10.1093/bioinformatics/btw411

Banga, J. R. 2008. Optimization in computational systems biology. BMC Syst. Biol. 2:47. doi:10.1186/1752-0509-2-47

Barreiro, X. R., and A. F. Villaverde. 2022. Benchmarking tools for a priori identifiability analysis. Preprint. October 17, 2022. [accessed November 23, 2022]. doi:10.48550/arxiv.2207.09745. https://doi.org/10.1093/bioinformatics/btad065

Batstone, D. J., J. Keller, I. Angelidaki, S. V Kalyuzhnyi, S. G. Pavlostathis, A. Rozzi, W. T. Sanders, H. Siegrist, and V. A. Vavilin. 2002. The IWA anaerobic digestion model no 1 (ADM1). Water Sci. Technol. 45:65–73.

Bellu, G., M. P. Saccomani, S. Audoly, and L. D'Angio. 2007. DAISY: A new software tool to test global identifiability of biological and physiological systems. Comput. Methods Programs Biomed. 88:52–61. doi:10.1016/j.cmpb.2007.07.002

Bezanson, J., A. Edelman, S. Karpinski, and V. B. Shah. 2017. Julia: a fresh approach to numerical computing. SIAM Rev. 59:65–98. doi:10.1137/141000671

Boer, H. M., S. T. Butler, C. Stotzel, M. F. Te Pas, R. F. Veerkamp, and H. Woelders. 2017. Validation of a mathematical model of the bovine estrous cycle for cows with different estrous cycle characteristics. Animal. 11:1991–2001. doi:10.1017/S175173111700026X

Boston, R. C., P. Wilkins, and L. O. Tedeschi. 2007. Identifiability and accuracy: two critical problems associated with the application of models in nutrition and the health sciences. In: Mathematical modeling for nutrition and health sciences. Roanoke, VA: University of Pennsylvania; p 161–193.

Cantet, R. J. C., and E. P. Cappa. 2008. On identifiability of (co)variance components in animal models with competition effects. J. Anim. Breed. Genet. 125:371–381. doi:10.1111/j.1439-0388.2008.00743.x

Chiş, O., J. R. Banga, and E. Balsa-Canto. 2011. GenSSI: a software toolbox for structural identifiability analysis of biological models. Bioinformatics. 27:2610–2611. doi:10.1093/bioinformatics/btr431

Chis, O. T., J. R. Banga, and E. Balsa-Canto. 2011. Structural identifiability of systems biology models: a critical comparison of methods. PLoS One. 6:e27755. doi:10.1371/journal.pone.0027755

Denis-Vidal, L., G. Joly-Blanchard, and C. Noiret. 2001. Some effective approaches to check the identifiability of uncontrolled nonlinear systems. Math. Comput. Simul. 57:35–44. doi:10.1016/s0378-4754(01)00274-9

Desmond-Le Quemener, E., and T. Bouchez. 2014. A thermodynamic theory of microbial growth. ISME J. 8:1747–1751. doi:10.1038/ismej.2014.7

Díaz-Seoane, S., X. Rey-Barreiro, and A. F. Villaverde. 2023. STRIKE-GOLDD 4.0: user-friendly, efficient analysis of structural identifiability and observability. Bioinformatics. 39:btac748. doi:10.1093/BIOINFORMATICS/BTAC748

Doeschl-Wilson, A. B., W. Brindle, G. Emmans, and I. Kyriazakis. 2009. Unravelling the relationship between animal growth and immune response during micro-parasitic infections. PLoS One. 4:e7508. doi:10.1371/journal.pone.0007508

Doeschl-Wilson, A., A. Wilson, J. Nielsen, H. Nauwynck, A. Archibald, and T. Ait-Ali. 2016. Combining laboratory and mathematical models to infer mechanisms underlying kinetic changes in macrophage susceptibility to an RNA virus. BMC Syst. Biol. 10:101. doi:10.1186/s12918-016-0345-5

Dong, R., C. Goodbrake, H. A. Harrington, and G. Pogudin. 2022. Differential elimination for dynamical models via projections with applications to structural identifiability. Preprint. November 23, 2022. [accessed November 25. 2022]. doi:10.48550/arxiv.2111.00991. http://doi.org/10.1137/22M1469067

Dougherty, H. C., E. Kebreab, M. Evered, B. A. Little, A. B. Ingham, J. V. Nolan, R. S. Hegarty, D. Pacheco, and M. J. Mcphee. 2017. The AusBeef model for beef production: II. Sensitivity analysis. J. Agric. Sci. 155:1459–1474. doi:10.1017/s0021859617000430

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

Ilmer, I., A. Ovchinnikov, and G. Pogudin. 2021. Web-Based Structural Identifiability Analyzer. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 12881 LNBI:254–265. doi:10.1007/978-3-030-85633-5_17

Joubert, D., J. D. Stigter, and J. Molenaar. 2020. An efficient procedure to assist in the re-parametrization of structurally unidentifiable models. Math. Biosci. 323:108328. doi:10.1016/j.mbs.2020.108328

Joubert, D., J. D. Stigter, and J. Molenaar. 2021. Assessing the role of initial conditions in the local structural identifiability of large dynamic models. Sci. Reports 11:1–15. doi:10.1038/s41598-021-96293-9

Karlsson, J., M. Anguelova, and M. Jirstrand. 2012. An Efficient Method for Structural Identifiability Analysis of Large Dynamic Systems. IFAC Proc. 45:941–946. doi:10.3182/20120711-3-BE-2027.00381

Kettle, H., G. Holtrop, P. Louis, and H. J. Flint. 2018. microPop: Modelling microbial populations and communities in R. Methods Ecol. Evol. 9:399–409. doi:10.1111/2041-210x.12873

Lam, N. N., P. D. Docherty, and R. Murray. 2022. Practical identifiability of parametrised models: A review of benefits and limitations of various approaches. Math. Comput. Simul. 199:202–216. doi:10.1016/j.matcom.2022.03.020

Lema-Perez, L., R. Muñoz-Tamayo, J. Garcia-Tirado, and H. Alvarez. 2019. On parameter interpretability of phenomenological-based semiphysical models in biology. Informatics Med. Unlocked. 15:100158. doi:10.1016/j.imu.2019.02.002

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

van Lingen, H. J., J. G. Fadel, L. E. Moraes, A. Bannink, and J. Dijkstra. 2019. Bayesian mechanistic modeling of thermodynamically controlled volatile fatty acid, hydrogen and methane production in the bovine rumen. J. Theor. Biol. 480:150–165. doi:10.1016/j.jtbi.2019.08.008

Macé, T., E. González-García, G. Kövér, D. Hazard, and M. Taghipoor. 2020. PhenoBR: a model to phenotype body condition dynamics in meat sheep. Preprint. December 2, 2020. [accessed November 23, 2022]. doi:10.1101/2020.12.01.407098.

Massonis, G., J. R. Banga, and A. F. Villaverde. 2021. AutoRepar: A method to obtain identifiable and observable reparameterizations of dynamic models with mechanistic insights. Int. J. Robust Nonlinear Control. 33:5039–5057. doi:10.1002/rnc.5887

Meshkat, N., C. Er-zhen Kuo, and J. DiStefano. 2014. On Finding and Using Identifiable Parameter Combinations in Nonlinear Dynamic Systems Biology Models and COMBOS: A Novel Web Implementation. PLoS One. 9:e110261. doi:10.1371/JOURNAL.PONE.0110261

Miao, H. Y., X. H. Xia, A. S. Perelson, and H. L. Wu. 2011. On Identifiability of Nonlinear ODE Models and Applications in Viral Dynamics. SIAM Rev. 53:3–39. doi:10.1137/090757009

Muñoz-Tamayo, R., J. C. Chagas, M. Ramin, and S. J. Krizsan. 2021. Modelling the impact of the macroalgae Asparagopsis taxiformis on rumen microbial fermentation and methane production. Peer Community J. 1:e7. doi:10.24072/PCJOURNAL.11

Muñoz-Tamayo, R., S. Giger-Reverdin, and D. Sauvant. 2016. Mechanistic modelling of in vitro fermentation and methane production by rumen microbiota. Anim. Feed Sci. Technol. 220:1–21. doi:10.1016/j.anifeedsci.2016.07.005

Muñoz-Tamayo, R., J. de Groot, E. Bakx, P. A. Wierenga, H. Gruppen, M. H. Zwietering, and L. Sijtsma. 2011. Hydrolysis of β-casein by the cell-envelope-located PI-type protease of Lactococcus lactis: A modelling approach. Int. Dairy J. 21:755–762. doi:10.1016/j.idairyj.2011.03.012

Muñoz-Tamayo, R., B. Laroche, M. Leclerc, and E. Walter. 2009. IDEAS: A parameter identification toolbox with symbolic analysis of uncertainty and its application to biological modelling. In: IFAC Proceedings Volumes. Vol. 42. p. 1271–1276.

Muñoz-Tamayo, R., B. L. Nielsen, M. Gagaoua, F. Gondret, E. T. Krause, D. P. Morgavi, I. A. S. Olsson, M. Pastell, M. Taghipoor, L. Tedeschi, et al. 2022. Seven steps to enhance open science practices in animal science. PNAS Nexus. 1:pgac106. doi:10.1093/pnasnexus/pgac106

Muñoz-Tamayo, R., M. Popova, M. Tillier, D. P. Morgavi, J. P. Morel, G. Fonty, and N. Morel-Desrosiers. 2019a. Hydrogenotrophic methanogens of the mammalian gut: Functionally similar, thermodynamically different—a modelling approach. PLoS One. 14:e0226243. doi:10.1371/journal.pone.0226243

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

Muñoz-Tamayo, R., J. F. Ramírez Agudelo, R. J. Dewhurst, G. Miller, T. Vernon, and H. Kettle. 2019b. A parsimonious software sensor for estimating the individual dynamic pattern of methane emissions from cattle. Animal. 13:1180–1187. doi:10.1017/S1751731118002550

Muñoz-Tamayo, R., and L. O. Tedeschi. 2022. The power of identifiability analysis for dynamic modeling in animal science: A practitioner approach. Preprint. November 29, 2022 – [accessed November 30, 2022]. doi:10.5281/ZENODO.7378158

Muñoz-Tamayo, R. 2023. Examples of identifiability analysis testing. doi:10.5281/ZENODO.7563737

Murphy, S. A., and A. W. Van Der Vaart. 2000. On profile likelihood. J. Am. Stat. Assoc. 95:449–465. doi:10.1080/01621459.2000.10474219

Nguyen-Ba, H., J. van Milgen, and M. Taghipoor. 2020. A procedure to quantify the feed intake response of growing pigs to perturbations. Animal. 14:253–260. doi:10.1017/S1751731119001976

Nimmegeers, P., J. Lauwers, D. Telen, F. Logist, and J. Van Impe. 2017. Identifiability of large-scale nonlinear dynamic network models applied to the ADM1-case study. Math. Biosci. 288:21–34. doi:10.1016/j.mbs.2017.02.008

Ovchinnikov, A., A. Pillay, G. Pogudin, and T. Scanlon. 2021. Computing all identifiable functions of parameters for ODE models. http://arxiv.org/abs/2004.07774

Poppi, D. P., W. C. Ellis, J. H. Matis, and C. E. Lascano. 2001. Marker concentration patterns of labelled leaf and stem particles in the rumen of cattle grazing bermuda grass (Cynodon dactylon) analysed by reference to a raft model*. Br. J. Nutr. 85:553–563. doi:10.1079/bjn2001317

Puillet, L., O. Martin, M. Tichit, and D. Sauvant. 2008. Simple representation of physiological regulations in a model of lactating female: Application to the dairy goat. Animal. 2:235–246. doi:10.1017/S1751731107001140

Puillet, L., D. Réale, and N. C. Friggens. 2016. Disentangling the relative roles of resource acquisition and allocation on animal feed efficiency: Insights from a dairy cow model. Genet. Sel. Evol. 48:1–16. doi:10.1186/s12711-016-0251-8

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

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

Raue, A., B. Steiert, M. Schelker, C. Kreutz, T. Maiwald, H. Hass, J. Vanlier, C. Tönsing, L. Adlung, R. Engesser, et al. 2015. Data2Dynamics: a modeling environment tailored to parameter estimation in dynamical systems. Bioinformatics. 31:3558–3560. doi:10.1093/bioinformatics/btv405

Reed, K. F., G. B. Arhonditsis, J. France, and E. Kebreab. 2016. Technical note: Bayesian calibration of dynamic ruminant nutrition models. J. Dairy Sci. 99:6362–6370. doi:10.3168/jds.2015-10708

Revilla, M., N. C. Friggens, L. P. Broudiscou, G. Lemonnier, F. Blanc, L. Ravon, M. J. Mercat, Y. Billon, C. Rogel-Gaillard, N. Le Floch, et al. 2019. Towards the quantitative characterisation of piglets' robustness to weaning: a modelling approach. Animal. 13:2536–2546. doi:10.1017/S1751731119000843

Roper, R. T., M. P. Saccomani, and P. Vicini. 2010. Cellular signaling identifiability analysis: A case study. J. Theor. Biol. 264:528–537. doi:10.1016/j.jtbi.2010.02.029

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

Saccomani, M. P., G. Bellu, S. Audoly, and L. D'Angió. 2019. A new version of DAISY to test structural identifiability of biological models. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 11773 LNBI:329–334. doi:10.1007/978-3-030-31304-3

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

Saltelli, A., K. Chan, and E. M. Scott. 2000. Sensitivity analysis. Wiley series in probability and statistics. Chichester: John Wiley & Sons.

Sedoglavic, A. 2002. A Probabilistic Algorithm to Test Local Algebraic Observability in Polynomial Time. J. Symb. Comput. 33:735–755. doi:10.1006/jsco.2002.0532

Shariati, M. M., I. R. Korsgaard, and D. Sorensen. 2009. Identifiability of parameters and behaviour of MCMC chains: a case study using the reaction norm model. J. Anim. Breed. Genet. 126:92–102. doi:10.1111/j.1439-0388.2008.00773.x

Shi, X., and M. N. Chatzis. 2022. An efficient algorithm to test the observability of rational nonlinear systems with unmeasured inputs. Mech. Syst. Signal Process. 165:108345. doi:10.1016/j.ymssp.2021.108345

Shorten, P. R., A. M. Ledgard, M. Donnison, P. L. Pfeffer, R. M. McDonald, and D. K. Berg. 2018. A mathematical model of the interaction between bovine blastocyst developmental stage and progesterone-stimulated uterine factors on differential embryonic development observed on Day 15 of gestation. J. Dairy Sci. 101:736–751. doi:10.3168/jds.2017-12845

Soetaert, K., and T. Petzoldt. 2010. Inverse Modelling, Sensitivity and Monte Carlo Analysis in R Using Package FME. J. Stat. Softw. 33:1–28. doi:10.18637/JSS.V033.I03

Stigter, J. D., and D. Joubert. 2021. Computing Measures of Identifiability, Observability, and Controllability for a Dynamic System Model with the StrucID App. IFAC-PapersOnLine. 54:138–143. doi:10.1016/j.ifacol.2021.08.34

Tedeschi, L. O., and R. Boston. 2010. Identifiability and accuracy: a closer look at contemporary contributions and changes in these vital areas of mathematical modelling. Model. Nutr. Dig. Util. Farm Anim. 91–99.

Tedeschi, L. O., and D. G. Fox. 2009. Predicting milk and forage intake of nursing calves. J. Anim. Sci. 87:3380–3391. doi:10.2527/jas.2009-2014

Tedeschi, L. O. 2006. Assessment of the adequacy of mathematical models. Agric. Syst. 89:225–247. doi:10.1016/j.agsy.2005.11.004

Vargas-Villamil, L. M., L. O. Tedeschi, S. Medina-Peralta, F. Izquierdo-Reyes, J. Navarro-Alberto, and R. González-Garduño. 2020. A multi-inverse approach for a holistic understanding of applied animal science systems. Animal. 14:s238–s249. doi:10.1017/S1751731120000877

Villaverde, A. F., A. Barreiro, and A. Papachristodoulou. 2016. Structural identifiability of dynamic systems biology models. PLoS Comput. Biol. 12:e1005153–e1005122. doi:10.1371/journal.pcbi.1005153

Villaverde, A. F., E. Raimundez, J. Hasenauer, and J. R. Banga. 2022. Assessment of Prediction Uncertainty Quantification Methods in Systems Biology. IEEE/ACM Trans. Comput. Biol. Bioinforma. 20:1725–1736. doi:10.1109/tcbb.2022.3213914

Walter, E., and L. Pronzato. 1996. On the identifiability and distinguishability of nonlinear parametric models. Math. Comput. Simul. 42:125–134. doi:10.1016/0378-4754(95)00123-9

Walter, E., and L. Pronzato. 1997. Identification of Parametric Models from Experimental Data London: Springer.

White, L. J., N. D. Evans, T. J. G. M. Lam, Y. H. Schukken, G. F. Medley, K. R. Godfrey, and M. J. Chappell. 2002. The structural identifiability and parameter estimation of a multispecies model for the transmission of mastitis in dairy cows with postmilking teat disinfection. Math. Biosci. 180:275–291. doi:10.1016/s0025-5564(02)00130-x

Wieland, F. G., A. L. Hauber, M. Rosenblatt, C. Tönsing, and J. Timmer. 2021. On structural and practical identifiability. Curr. Opin. Syst. Biol 25:60–69. doi:10.1016/j.coisb.2021.03.005