



HAL
open science

Observability, Identifiability and Epidemiology A survey

Frédéric Hamelin, Abderrahman Iggidr, Alain Rapaport, Gauthier Sallet

► **To cite this version:**

Frédéric Hamelin, Abderrahman Iggidr, Alain Rapaport, Gauthier Sallet. Observability, Identifiability and Epidemiology A survey. 2020. hal-02995562v1

HAL Id: hal-02995562

<https://hal.science/hal-02995562v1>

Preprint submitted on 20 Nov 2020 (v1), last revised 25 May 2023 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Observability, Identifiability and Epidemiology

A survey

F. Hamelin², A. Iggidr¹, A. Rapaport⁴, G. Sallet³

November 20, 2020

¹INRIA

² Agro Campus Ouest Rennes

³ IECL, Université de Lorraine

⁴ INRAE

Abstract

In this document we introduce the concepts of Observability and Identifiability in Mathematical Epidemiology. We show that, even for simple and well known models, these properties are not always fulfilled. We also consider the problem of practical observability and identifiability which are connected to sensitivity and numerical condition numbers.

1 Introduction

1.1 Prologue

Many papers in Epidemiology have the following structure :

- A model is proposed,
- some parameters are given, extracted from literature,
- then the remaining unknown parameters are estimated by fitting the model to some observed data.

Fitting is done usually by using an optimisation algorithm and use for example a least square method or a maximum likelihood estimation.

To validate the parameters estimation, the algorithm is used to recover the parameters from noisy simulated data obtained from the model with some values given to the parameters.

The objective of this paper is to show that this procedure does not prove anything and that an examination of the identifiability of parameters is a prerequisite before a numerical determination of parameters. We will review different methods to study identifiability and observability and then consider the problem of numerical identifiability. Our touchstone will be the most famous, however simple, model in Mathematical Epidemiology, SIR model of Kermack and Mckendrick [44]

Parameter identifiability analysis addresses the problem of which unknown parameters of an ODE model can uniquely recovered from observed data. We will show that, even for very simple models, identifiability is not automatic.

The problem of identifiability for Epidemiological models is rarely addressed:

A research in MathScinet with

- EPID* AND IDENTIFIABILITY gives only 4 papers,
- biolog* AND epid* shows 30 papers,
- on the other hand epidem* AND parameter returns 68 publications.

Only a small part of all these publications address the problem of identifiability. Usually the paper proposes an algorithm and very often the identifiability is not tackled.

The following publications consider the problem of identifiability in Epidemiological models. However the majority is published outside biomathematics journals.

[4, 5, 11, 17, 16, 28, 29, 39, 48, 56, 58, 60, 65, 64, 75, 76, 77, 78, 82, 83]

1.2 Definitions

The question of parameter identifiability originates from control theory and is related to observability and controllability [69]. The first appearance is in Kalman [42] and is now sixty years old. Identifiability is related to observability : observability of a model is the ability to reconstruct the state of a system from the observation. In the language of dynamical systems with inputs and outputs, which is the standard paradigm in control systems theory,

an input-output relation is defined. The inputs, are also called the control, are considered as known. We will only consider system without control, which is a peculiar case where the control is a singleton. When controls are known, with more information observability/identifiability is sometimes more easy. These problems has been rarely considered for uncontrolled systems whereas many methods have been developed for controlled systems. To be more precise, let consider the following system in \mathbb{R}^n

$$\Sigma \begin{cases} \dot{x}(t) &= f(x(t)) \\ x(0) &= x_0 \\ y(t) &= h(x(t)), \end{cases} \quad (1)$$

where we have denoted $\dot{x}(t) = \frac{d}{dt}x$.

The ODE $\dot{x} = f(x)$ is the dynamics of the system and x is called the state of the system. To avoid technical details we will assume that for any initial condition x_0 , there exists an unique solution denoted $x(t, x_0)$ such that $x(0, x_0) = x_0$ and

$$\frac{d}{dt}x(t, x_0) = f(x(t, x_0)).$$

We will assume that this solution $x(t, x_0)$ is defined for any time $t \geq 0$, which is often the case in epidemiological models. This often the case with epidemiological models for which state space is a positively compact invariant set. Then we will assume that the system is defined on a positively invariant compact set Ω , which means that any solution starting in Ω stays in the compact and this implies that the solution is defined for ant $t \geq 0$. This situation is also often encountered in biological systems.

The output (or observation) of the system is given by $h(x)$ where h is a differentiable function $h : x \in \mathbb{R}^n \mapsto h(x) \in \mathbb{R}^m$. The set \mathbb{R}^m is the observation space. We will denote by $h(t, x_0)$ or $y(t, x_0)$ the observation at time t .

Definition 1.1 (Observability)

The system (1) is observable if for two distincts initial condition $x_1 \neq x_2$ there exists a time $t \geq 0$ such that

$$h(x(t, x_1)) \neq h(x(t, x_2))$$

Two states are called indistinguishable if for any $t \geq 0$ we have

$$h(x(t, x_1)) = h(x(t, x_2)).$$

Indistinguishability means that it is impossible to differentiate the evolution of the system, from to distincts initial conditions, by considering only with the observation.

Now we consider a system depending on a parameter $\theta \in \mathbb{R}^p$

$$\begin{cases} \dot{x}(t) &= f(x(t), \theta) \\ x(0) &= x_0 \\ y(t) &= h(x(t), \theta). \end{cases} \quad (2)$$

Identifiability is the ability to recover the unknown parameters from the observation. We denote by $x(t, x_0, \theta)$ the solution of (2) for an intiiial condition x_0 .

Definition 1.2

System (2) is sais to be identifiable if for any distincts parameters $\theta_1 \neq \theta_2$, there exists $t \geq 0$ such that

$$h(x(t, x_0, \theta_1)) \neq x(t, x_0, \theta_2).$$

There is an obvious similarity between observability and identifiability. Actually we will say that (2) is observable and identifiable if the augmented system

$$\begin{cases} \dot{x}(t) &= f(x(t), \theta) \\ \dot{\theta} &= 0 \\ y(t) &= h(x(t), \theta). \end{cases} \quad (3)$$

is observable.

Actually for an epidemiological model it is unlikely to know the initial conditions and it has long been recognized that initial conditions play a role in identifying the parameters [26, 50, 74, 83, 66].

What we have called identifiability is also known as structural identifiability. This expression has been coined by Bellman and K.J. Åström [10] in 1970. This is to stress that identifiability depends only on the dynamics and the observation, under ideal conditions of noise-free observations and error-free model. This is a mathematical and *a priori* problem [39].

1.3 History

The observability concept has been introduced by Kalman [42] in 1960 for linear systems.

For nonlinear systems, observability has been characterized circa 1970 [35, 37]. The definition is given by Hermann and Krener in the framework of differential geometry.

Identifiability and structural identifiability was introduced in compartmental analysis in 1970 by Richard Bellman and K.J. Åström [10] in a paper in Mathematical Biosciences.

The problem of identifiability is now addressed in books [49, 81, 80, 79]. Numerical identifiability of linear control system is implemented in MATLAB, SCILAB.

Identifiability of nonlinear systems has been addressed in different context and the first systematic approach is by Tunali and Tarn [74] in the differential geometry framework.

The introduction of the concepts of differential algebra in control theory is due to Michel Fliess around 1990 [26, 27, 31] followed by Torkel Glad [33, 50].

Identifiability is a general problem which has received different names depending on the community :

- observation, identification;
- data assimilation ;
- inverse problems;
- parameters estimation.

Data assimilation is used in en meteorology and oceanography. [46, 72].

A direct problem is considering a model which, when introducing an input, gives an observed output. The parameters are considered as known. Conversely the inverse problem is to reconstruct the parameters from the knowledge of the output [73].

Finally expression of parameters estimation is currently used in the probability and statistics domain.

1.4 Identifiability in Mathematical Epidemiology

Identifiability is well known in Biomathematics from 1970, as we have already said, with the paper of Bellman and Åström.

In the other hand, considering identifiability in Mathematical Epididemiolgy is relatively recent [75, 60, 83, 65, 56, 28, 29].

The first paper, to our knowledge, considering identifiability of an intra-host model of HIV is by X. Xia et C.F. Moog [83], and has been published in 2003 in the control journal *IEEE Trans. Automat. Control*.

2 Observability and Identifiability

We consider the following system (1)

$$\Sigma \begin{cases} \dot{x} &= f(x) \\ y &= h(x), \end{cases} \quad (4)$$

Where we assume again to avoid technical details that f and g are \mathcal{C}^∞ functions. The function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is called a vector field.

The classical definition of Lie derivative of a \mathcal{C}^∞ function $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ with respect to the vector field f is given by

$$\mathcal{L}_f(g) = \left. \frac{d}{dt} g(x(t, t_0, x_0)) \right|_{t=0} = \langle \nabla g(x) | f(x) \rangle,$$

with ∇g is the gradient of g and $\langle | \rangle$ the inner product of \mathbb{R}^n .

2.1 Observability

2.1.1 Observability with Differential Geometry

We denote by $h = (h_1, \dots, h_m)$ the components of the observation function h .

Each function h_i is a function \mathcal{C}^∞ from the state space \mathbb{R}^n to \mathbb{R} .

Definition 2.1 ([37])

The observation space \mathcal{O} of (4) is the subspace of the vector space $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$ containing h_i and invariant by the Lie derivative \mathcal{L}_f of the dynamics

The observation space is generated by the different Lie derivatives of the h_i :

$$\mathcal{O} = \{h_i, \dots, \mathcal{L}_f h_i, \dots, \mathcal{L}_f^k h_i, \dots\}_{\{\text{span}, i=1, \dots, m\}}$$

We have the following result

Theorem 2.1 For an analytic system (i.e., f and h are analytics functions) the observability is equivalent to the separation of the points of the state space \mathbb{R}^n by \mathcal{O} i.e., if $x_1 \neq x_2$ there exists $g \in \mathcal{O}$ such that $g(x_1) \neq g(x_2)$.

Proof

By analyticity we have

$$y(t, x_0) = h(x(t, x_0)) = \sum_{n \geq 0} \frac{t^n}{n!} \frac{d^n}{dt^n} h(x(t, x_0)) \Big|_{t=0}.$$

but, by induction we have the following relation

$$\frac{d^n}{dt^n} h(x(t, x_0)) \Big|_{t=0} = \mathcal{L}_f^n h(x_0).$$

Then a necessary and sufficient condition to distinguish $x_1 \neq x_2$ is that there exists n such that

$$\mathcal{L}_f^n h(x_1) \neq \mathcal{L}_f^n h(x_2).$$

■

We have defined a global observability concept, but it might be necessary to travel a considerable distance or for a long time to distinguish between points of \mathbb{R}^n . Therefore a local concept of observability is introduced [37].

Definition 2.2

The system (4) Σ is said locally observable if, for any x_0 , for any open set U containing x_0 , x_0 is distinguishable from all the points of U for the restricted system $\Sigma|U$.

The system (4) Σ is locally weakly observable if for any x_0 there exists an open set U containing x_0 , such that for any neighborhood V with $x_0 \in V \subset U$, x_0 is distinguishable for $\Sigma|V$ from all the points of V .

Intuitively a system is locally weakly observable if one can instantaneously distinguish each point from its neighbors.

The local weak observability can be characterized

Definition 2.3 [71]

We define

$$d\mathcal{O} = \{d\psi \mid \psi \in \mathcal{O}\}$$

Definition 2.4

A system Σ is said to satisfy the observability rank condition in x if the dimension of $d\mathcal{O}$ in x satisfies

$$\dim(d\mathcal{O}(x)) = n$$

$d\mathcal{O}$ is generated by the gradients of the $\mathcal{L}_f^k h$.

Theorem 2.2 (Hermann-Krener [37])

If Σ satisfies the observability rank condition (ORC) at x_0 then Σ is locally weakly observable at x_0 .

Proof

Since $\dim(d\mathcal{O}(x)) = n$, there exists n functions $\varphi_1, \dots, \varphi_n \in \mathcal{O}$ such that the gradients $d\varphi_1, \dots, d\varphi_n$ are linearly independent.

Therefore the function $\Phi : x \mapsto (\varphi_1(x), \dots, \varphi_n(x))$ has a nonsingular Jacobian in x_0 .

As a consequence there exists an open set U containing x_0 where Ψ is a bijection . (t)=

On any open set $V \subset U$ assume that we have $h(x(t, x_0)) = h(x(t, x_1))$ for $x_0 \neq x_1$ and $t > 0$ sufficiently small. Then by derivation we have $\mathcal{L}_f^k(x_0) = \mathcal{L}_f^k(x_1)$ for any k therefore $\varphi_i(x_0) = \varphi_i(x_1)$ which is a contradiction. ■

Proposition 2.1

For an analytic system if the observability rank condition is satisfied everywhere the system is locally observable, hence observable

Proof

This is due to the fact that $h(x(t, x_0))$ is sum of his Taylor sequence . The rank condition implies, by the same reason as before, that the coefficients of the Taylor sequence separates points. ■

Exemple 2.1

We consider the SIR model of Kermack-McKendrick [44] for which the parameters β, γ are assumed to be known :

$$\left\{ \begin{array}{l} \dot{S} = -\beta \frac{S}{N} I \\ \dot{I} = \beta \frac{S}{N} I - \gamma I \\ \dot{R} = \gamma I \\ y = \gamma I, \end{array} \right. \quad (5)$$

with $N = S + I + R$ We have $\dot{y} = \gamma (\beta \frac{SI}{N} - \gamma I)$. A generating set of \mathcal{O} contain I and $\beta \frac{SI}{N}$. These two functions, when $I \neq 0$, separates points (using $R = N - S - I$). The system is observable on $\mathbb{R}^3 \setminus \mathbb{R} \times \{0\}$.

With the gradients the matrix

$$\begin{bmatrix} 0 & \beta \frac{I}{N} \\ 1 & \frac{S}{N} \end{bmatrix}$$

is non singular if $I \neq 0$. The rank condition is satisfied on $\mathbb{R}^2 \setminus \mathbb{R} \times \{0\}$.

2.1.2 Observability and Differential algebra

From 1970 differential geometry was the tool for studying nonlinear systems in control theory. Circa 1985 Michel Fliess propose to use Differential Algebra for the analysis of nonlinear systems.

Intuitively observability for system (4) is that the state x can be expressed as an algebraic expression of y and his derivatives $y^{(k)}$ [27, 26].

Definition 2.5 [27, 26]

A system is said to be algebraically observable if the states can be reconstructed by solving algebraic equations depending only of the observation and his derivatives.

Note that the systems under consideration are rational systems, i.e., the functions f and h are rational functions [26, 27]. A more precise definition can be given using the formalism of Differential Algebra. Differential Algebra can be considered as a generalization to differential equations of the concepts of commutative algebra and algebraic geometry.

This theory, founded by Ritt, is an appropriate framework fo definition of algebraic observability. We will not go in further details and the interested reader can consult Michel Fliess publications [30, 32].

Exemple 2.2

Consider the SIR model (5).

This system evolves in the positively invariant set (which makes biological sense)

$$K = \{(S, I, R) \mid S > 0, I > 0, N = S + I + R\}$$

which implies $y \neq 0$.

Then we have the relations (we can divide by y)

$$\frac{S}{N} = \frac{1}{\beta} \frac{\dot{y} + \gamma y}{y}, \quad \frac{S}{N^2} = \frac{1}{\beta^2} \frac{\dot{y}^2 - y \ddot{y}}{y^3}.$$

We recover N as a rational expression of y, \dot{y}, \ddot{y} , then also for S and R . The system is algebraically observable in K .

2.1.3 A Link between the two definitions

Consider a state-output system

$$\Sigma \begin{cases} \dot{x} = f(x) \\ y = h(x), \end{cases} \quad (6)$$

where f and h are polynomial or rational functions in x .

We have two rational relations $x - f(x) = 0$ and $y - h(x) = 0$. The last relation can be derived $\dot{y} - \frac{\partial}{\partial x} h \dot{x} = \dot{y} - \frac{\partial}{\partial x} (h) f(x) = 0$. By induction we will obtain a sequence of rational relations $Q_j(x, y) = 0$.

It can be shown [4, 27, 68] that to obtain algebraic observability it is sufficient to consider the matrix

$$\begin{bmatrix} h_1 & \frac{\partial}{\partial x_1} \mathcal{L}_f(h_1) & \cdots & \frac{\partial}{\partial x_n} \mathcal{L}_f(h_1) \\ \vdots & \vdots & \ddots & \vdots \\ h_n & \frac{\partial}{\partial x_1} \mathcal{L}_f^{n-1}(h_1) & \cdots & \frac{\partial}{\partial x_n} \mathcal{L}_f^{n-1}(h_1) \end{bmatrix}$$

and to check that his rank is n the dimension of the state space.

This is the Herman-Krener criterion for local weak observability. But the system is analytic.

Is the number of Lie derivative bounded ? It is proved that it is sufficient compute no more than $n - 1$ derivatives [4, 27, 68].

Theorem 2.3 [27, 71]

Consider (6) where f and g are rational functions. To check the observability rank condition ($n - 1$) Lie derivatives are sufficient.

Proof

A detailed proof is given in cited references. However an outline goes like this :

Assume that k is the first integer such that $\{dh, d\mathcal{L}_f h, \dots, d\mathcal{L}_f^k h\}$ are linearly dependent with rational coefficients

Then there exists $k + 1$ functions, not all of them zero, $g_i \in \mathbb{R}_x$ such that

$$\sum_{i=0}^k g_i \cdot d\mathcal{L}_f^i h = 0$$

Particularly, with k definition, we have $g_k \neq 0$. We will apply \mathcal{L}_f to this relation

$$\sum_{i=0}^k \mathcal{L}_f g_i \cdot d\mathcal{L}_f^i h + g_i \cdot \mathcal{L}_f d\mathcal{L}_f^i h = 0.$$

It is well known that d and \mathcal{L}_f commute. Therefore $\mathcal{L}_f d\mathcal{L}_f^i h = d\mathcal{L}_f^{i+1} h$. Because the system is rational $\mathcal{L}_f g_i \in \mathbb{R}_x$. Since $g_k \neq 0$ this prove that $d\mathcal{L}_f^{k+1}$ is a linear combination of the previous Lie derivatives

This ends the proof. ■

The method of Differential Algebra promoted by Fliess et al. has the advantage to be implemented inside a computer algebra software. Actually as soon as the system is in high dimension, the computations get rapidly involved. There exists now some software [68, 11, 17, 43].

2.2 Identifiability

It can happen that a system is identifiable and however not observable

Example 2.3

The following academic systems is identifiable and however not observable

$$\begin{cases} \dot{x}_1 = -\alpha(x_1 + x_2) \\ \dot{x}_2 = \alpha(x_1 - x_2) \\ y = \frac{1}{2}(x_1^2 + x_2^2) \end{cases} \quad (7)$$

We have $\dot{y} = -\alpha y$, $\ddot{y} = \alpha^2 y$ and $y^{(p)} = (-1)^p \alpha^p y$. The Jacobian

$$Jac [h, \mathcal{L}_f h, \mathcal{L}_f^2 h] = \begin{bmatrix} x_1 & -\alpha x_1 & \alpha^2 x_1 \\ x_2 & -\alpha x_2 & \alpha^2 x_2 \\ 0 & \alpha y & \alpha^2 y \end{bmatrix},$$

is clearly of rank 2.
 α is differentially algebraic on the field $\mathbb{R}\langle y \rangle$. Identifiability is in $\mathbb{R}^2 \setminus \{0\}$, this is immediate with $\alpha = -\dot{y}/y$.

2.2.1 Identifiability and observability

Since very often the initial conditions are not known, or partially known, we will consider in the remaining, the problem of identifiability and observability. Then we will consider the augmented system (3).

2.2.2 Characteristic set

We consider a polynomial system

$$\begin{cases} \dot{x}(t) = f(x(t), \theta) \\ x(0) = x_0 \\ y(t) = h(x(t), \theta), \end{cases} \quad (8)$$

with $x \in \mathbb{R}^n$, $\theta \in \mathbb{R}^p$ and $y \in \mathbb{R}^m$. We will consider differential polynomials, i.e., polynomial in $n + m$ variables and their derivatives, with coefficients in \mathbb{R} . For example $\dot{x} - f(x, \theta)$ and $\dot{y} - h(x, \theta)$ are differential polynomials in $\mathbb{R}\langle x, y \rangle$. We consider the parameter $\theta \in \mathbb{R}^p$ as a constant, i.e., $\dot{\theta} = 0$.

We have $n + m$ differential polynomial equations, n states and m observations. Intuitively we can obtain, by differentiating-multiplying by any differential polynomial-adding, an infinity of new equations. In other words the $n + m$ equations generate a differential ideal \mathcal{I} . Any x, y satisfying (6) will satisfy these equations in the ideal \mathcal{I} .

The idea behind the differential algebra techniques is to get a finite set of differential polynomials which describes and summarizes all the information in the generated ideal. Such a set \mathcal{C} is called a characteristic set. The details of the complete algorithm for constructing a characteristic set is rather involved and can be found in the references [45, 50, 62].

Among all the polynomials in \mathcal{I} we can consider the set \mathcal{I}^c of differential polynomials with only the observation y . Since $\mathcal{I}^c = \mathbb{R}[y] \cap \mathcal{I}$ this set is an ideal. It is possible to obtain a characteristic set for \mathcal{I}^c , namely $\mathcal{C} \cap \mathbb{R}[\theta]\langle x, y \rangle$. This set is obtained by eliminating the state variables from the equations [25].

Actually, since we have no input, this characteristic set is the differential polynomial in y with the lower order [41].

This set is also called, in the literature the input-output relation of the system [5, 11, 54, 40].

Making the polynomials monic in the input-output relations give a set of coefficients $c_i(\theta)$ and a map $c : \mathbb{R}^p \rightarrow \mathbb{R}^r$ which is called the exhaustive summary of the model [59, 64, 66, 5]. The injectivity of c from the parameter space is only a necessary condition for identifiability [66].

However the input-output relations do not depend of the initial conditions and since the identifiability is dependent of these initial conditions it can happen that, with c injective, the system is not identifiable.

Some authors [55, 75, 28] use the injectivity of c to ascertain the identifiability, which is false [66] section 3.3. A theorem is even given (Th 3.1 [28]). The following example shows the injectivity of c is not sufficient.

Exemple 2.4 (input-output relation is not sufficient)

consider the following compartmental system

$$\begin{cases} \dot{x}_1 &= -a_{21} x_1 + a_{12} x_2 \\ \dot{x}_2 &= -a_{12} x_2 \\ y &= x_1. \end{cases} \quad (9)$$

We have

$$\dot{y} = -a_{21} x_1 + a_{12} x_2 \quad \ddot{y} = -a_{21} \dot{y} - a_{12}^2 x_2 \quad \ddot{y} = -a_{21} \dot{y} - a_{12} (\dot{y} + a_{21} y)$$

The input-output relation is

$$\ddot{y} + (a_{12} + a_{21}) \dot{y} + a_{12} a_{21} y = 0.$$

Clearly the application $(a_{21}, a_{12}) \mapsto ((a_{12} + a_{21}, a_{12} a_{21}))$ is injective and however the system is not identifiable if $x_2 = 0$.

This can be seen in two ways :

- *it is easy to see that*

$$\begin{aligned}\ddot{y}^2 - \dot{y} \ddot{y} &= (\dot{y}^2 - y \ddot{y}) a_{21} a_{12}, \\ \dot{y} \ddot{y} - y y^{(3)} &= -(\dot{y}^2 - y \ddot{y}) (a_{12} + a_{21}).\end{aligned}$$

Then we need to have $(\dot{y}^2 - y \ddot{y}) \neq 0$ to recover the parameters.

$$(y \ddot{y} - \dot{y}^2) = a_{12} x_2 [(a_{21} - a_{12}) x_1 - a_{12} x_2].$$

Clearly we need $x_2 \neq 0$. We need also $a_{12} \neq 0$ for observability, since $a_{12} x_2 = \dot{y} + a_{21} x_1$.

- The second way is to compute the Jacobian of $[y, \dot{y}, \ddot{y}, y^{(3)}]$ for local identifiability and observability.

We have easily

$$\det [Jac [y, \dot{y}, \ddot{y}, y^{(3)}]] = -a_{12}^2 (a_{12} - a_{21}) [(a_{21} - a_{12}) x_1 - a_{12} x_2] x_2$$

2.3 How many measures to identify ?

In the case of analytic systems an answer has been given by par D. Aeyels [2, 1] and E. Sontag [70]:

For an analytic system with r parameters it is sufficient to randomly choose $2r + 1$ measures to distinguish two different states..

This means that generically (hence the term randomly) $2r + 1$ measures are sufficient.

3 The classical SIR model of Kermack-Mckendrick

The SIR model of Kermack et McKendrick [44] is certainly one of the most famous model in Epidemiology. It is given and studied in all of the classic books of Mathematical Epidemiology. This model appears in the book of Bailey, which is probably the first book in Mathematical Epidemiology. Some examples can be found in [3, 12, 13, 14, 21, 47, 52, 57].

The figure, in the original paper, fitting the model to plague data in Bombay during the 1906 year, is one of the most famous picture in Epidemiology. A research with SIR in MathScinet returns 11 106 articles.

In the quoted books the SIR model is fitted to data :

- In [12, 13, 14] the model is fitted with the Great Plague in Eyam on the year 1666;
- in [21] Influenza in England and Wales;
- in [47] a fitting is done with simulated noisy data;
- in [52, 12] where a chapter is devoted to fitting epidemiological models to data, a SIR model is used to represent an influenza outbreak in an English boarding school.

More recently two publications [6, 51] try to fit the Kermack-McKendrick model to the plague in Bombay.

As already said, before entering in algorithms to adjust parameters, an identifiability analysis must be done.

3.1 The different forms of the SIR model

The original model [44] is

$$\begin{cases} \dot{S} = -\tilde{\beta} S I \\ \dot{I} = \tilde{\beta} S I - \gamma I \\ \dot{R} = \gamma I \end{cases} \quad (10)$$

Where S, I, R represent respectively the number of susceptibles, infectious and removed individuals.

This model can also be found in another form

$$\begin{cases} \dot{S} = -\beta \frac{S}{N} I \\ \dot{I} = \beta \frac{S}{N} I - \gamma I \\ \dot{R} = \gamma I \end{cases} \quad (11)$$

Where $N = S + I + R$ is the total population.

One can pass from the first model to the other with $\tilde{\beta} = \frac{\beta}{N}$. This is not harmless since N , which is a parameter, is hidden in $\tilde{\beta}$.

For the basic reproduction ratio, N intervenes

$$\mathcal{R}_0 = \frac{\tilde{\beta} N}{\gamma} = \frac{\beta}{\gamma}$$

One of the greatest sources of confusion is the introduction of the term pseudo mass action by de Jong, Diekmann et Heesterbeek [23]. They correctly pointed out that $\beta S I$ represents the true mass action if S and I represent densities of hosts (numbers per unit area). This was the case for the original paper, for which the system was intended to model the plague in Bombay, which is an island. Several laboratory studies have found that the $\beta S I$ model is inadequate for describing pathogen [53]. The term $\beta \frac{S}{N} I$ is often called frequency-dependent or density dependent transmission.

3.2 Observability-Identifiability of the SIR model

It is surprising that the observability and identifiability of the original Kermack-Mckendrick model has not been more studied, since this system has been used many times to model an outbreak of infection. The observability and identifiability of the model SIR, with demography and constant population, has been first studied in 2005 [29] :

$$\begin{cases} \dot{S} &= \mu N - \beta \frac{S}{N} I - \mu S \\ \dot{I} &= \beta \frac{S}{N} I - (\gamma + \mu) I \\ \dot{R} &= (\gamma + \mu) I \end{cases} \quad (12)$$

This model has 7 parameters, including one for the observation $k I$. The result in [29] is that the system is neither observable nor identifiable.

In [76] the identifiability of (11) is addressed assuming that the initial conditions are known, and use only the input-output relation to conclude. This is incorrect, as we have seen, and moreover the input-output relation has

an error, as we will see later. With the initial conditions the identifiability is quite immediate if I or γI is observed, as we shall see but of dubious interest.

Consider the SIR model

$$\begin{cases} \dot{S} &= -\beta \frac{S}{N} I \\ \dot{I} &= \beta \frac{S}{N} I - \gamma I \\ y &= k I \end{cases} \quad (13)$$

The last equation has been omitted since $R = N - S - I$. Observation is $k I$, in other words only a percentage of infectious individuals are observed. This situation is used for example in [51, 63].

Theorem 3.1

The parameters of the system are $S, I, N, \beta, \gamma, k$.

The system (13) is neither observable, nor identifiable.

The functions $k I, \frac{\beta}{k N}, \frac{\beta S}{N}, \gamma$ are identifiable, equivalently $(k S, k I, \gamma, \frac{\beta}{k N})$ are identifiable.

Particularly if N is known and if $k = 1$ or $k = \gamma$ the system is identifiable and observable.

If $k = \gamma$ then $S, I, \gamma, \frac{\beta}{N}$ are identifiable.

Remarque 3.1

We could believe that if $k = \gamma$, with N unknown, then (10) is observable.

This is wrong because we reduce the system with $\dot{R} = \gamma I$. Certainly S, I are observable, but not R . Moreover \mathcal{R}_0 is not observable.

Proof

We will show that the parameters and the state can be expressed with the observation and their derivatives [27, 26]. Otherwise the identifiability will not be obtained

We consider the system on

$$\mathcal{I} = \{(s, I) \mid S > 0, I > 0, S + I < N\}$$

This open set \mathcal{I} is positively invariant. In any equilibrium point $(S_0, 0)$ the system is not observable. Therefore we assume $y \neq 0$ and also, for the same reason, $S \neq 0$.

We have (with $S \neq 0$ and $I \neq 0$)

$$\begin{aligned} y &= k I & \frac{\dot{y}}{y} &= \frac{\beta S}{N} - \gamma = h_1 & -\frac{\dot{h}_1}{y} &= \frac{\beta^2}{k N^2} S = h_2 \\ -\frac{\dot{h}_2}{y} &= \frac{\beta^3}{k^2 N^3} S = h_3 & \frac{h_3}{h_2} &= \frac{\beta}{k N} = \frac{\tilde{\beta}}{k} & \frac{h_2^2}{h_3} &= \frac{\beta S}{N} = \tilde{\beta} S \\ h_1 - \frac{h_2^2}{h_3} &= \gamma. \end{aligned}$$

All the derivative of y can be expressed as rational expressions of $k I, \frac{\beta}{k N}, \frac{\beta S}{N}, \gamma$.

Therefore the only informations obtained will be $k I, \frac{\beta}{k N}, \frac{\beta S}{N}, \gamma$ which are identifiable functions.

Now if N is known with $k = 1$ or $k = \gamma$, the parameters β, γ, S, I, R are polynomial functions of (y, \dot{y}, \ddot{y}) , hence the system is observable and identifiable.

Equivalently the variables $(k S, k I, \gamma, \frac{\beta}{k N})$ are identifiable. ■

3.2.1 Using input-output relations

We use elimination

$$\begin{aligned} \dot{y} &= \frac{\beta S}{N} y - \gamma y & \text{which gives} & \frac{\beta S}{N} = \frac{\dot{y} + \gamma y}{y} \\ \ddot{y} &= -\frac{\beta^2 S}{k N^2} y^2 + \frac{\beta S}{N} \dot{y} - \gamma \dot{y} & \text{which gives} & \ddot{y} = -\frac{\beta}{k N} y (\dot{y} + \gamma y) + \frac{\dot{y} + \gamma y}{y} \dot{y} - \gamma \dot{y} \end{aligned}$$

Equivalently

$$\dot{y} \ddot{y} + \frac{\beta}{k N} y^2 \dot{y} + \frac{\beta}{k N} \gamma y^3 - \dot{y}^2 = 0$$

We have already seen, with the notation in the preceding proof, that

$$\frac{\beta S}{N} = \gamma + h_1.$$

This shows that $k I, \frac{\beta}{k N}, \frac{\beta S}{N}, \gamma$ are identifiable.

3.2.2 Using ORC Herman-Krener condition

Consider the application $\Phi : (S, I, \beta, \gamma, k) \rightarrow (y, y^{(1)}, y^{(2)}, y^{(3)}, y^{(4)})$. We denote by $Jac \Phi = \frac{\partial \Phi}{\partial (S, I, \beta, \gamma, k)}$ the Jacobian of Φ . Then $\det(Jac \Phi) = 0$ which proves that the system is not identifiable neither observable.

On the other hand with $\Psi : (S, I, \beta, \gamma) \rightarrow (y, y^{(1)}, y^{(2)}, y^{(3)})$, denoting by $Jac \Psi = \frac{\partial \Psi}{\partial (S, I, \beta, \gamma)}$ the Jacobian of Ψ ,

$$\text{we have } \det(Jac \Psi) = -\frac{k^4 \beta^4 S I^6}{N^5} \neq 0.$$

This proves, with N and k known the local observability and identifiability.

3.2.3 SIR with cumulative incidence

Very often the observations are the cumulative numbers of infectious individuals. Is the SIR model still observable?

Then we consider the system where the observation is given by

$$y(t) = k \int_0^t \frac{S(\tau) I(\tau)}{N} d\tau. \quad (14)$$

This problem has been addressed for the SIR model with demography with constant population in [29]. Also is considered the identifiability with known initial conditions for (13) in [76] using input-output relations.

Theorem 3.2

The system (13) with cumulative incidence observation is neither observable, nor identifiable.

The parameters $(k S, k I, \gamma, \frac{\beta}{k N})$ are identifiable.

When N is known and when $k = 1$ or $k = \gamma$ the system is observable and identifiable.

Proof

A straightforward computation gives

$$\begin{aligned} \dot{y} &= k \frac{\beta S I}{N} & h_2 &= \frac{\ddot{y}}{\dot{y}} = \frac{\beta S}{N} - \frac{\beta I}{N} - \gamma \\ h_3 &= \frac{\dot{h}_2}{\dot{y}} = \frac{\beta}{N} \gamma I - 2 \frac{\beta}{k N} y & h_4 &= \dot{h}_3 = -\frac{\beta}{N} \gamma^2 I + \frac{\beta}{k N} \gamma y - 2 \frac{\beta}{k N} \dot{y} \end{aligned}$$

Finally

$$\frac{\beta}{k N} \gamma y + 2 \frac{\beta}{k N} \dot{y} + \gamma h_3 + h_4 = 0.$$

Differentiating this relation, if $\dot{y}^2 - y \ddot{y} \neq 0$, gives

$$2 \frac{\beta}{k N} = \frac{(\gamma \dot{h}_3 + \dot{h}_4) - (\gamma h_3 + h_4)}{\dot{y}^2 - y \ddot{y}}.$$

Since the zeroes of an analytic function are isolated we have to prove that $\dot{y}^2 - y \ddot{y} \neq 0$. If this is the case, this implies that $\frac{d}{dt} \left(\frac{\dot{y}}{y} \right) = 0$. Hence $y = C e^{\lambda t}$, therefore $h_2 = \lambda$, which implies $S - I = \text{constant}$, a contradiction. Since we have

$$k S - k I = \frac{\gamma + h_2}{\left(\frac{\beta}{k N} \right)} \quad k S \quad k I = \frac{\dot{y}}{\left(\frac{\beta}{k N} \right)}.$$

this proves our claim. ■

4 Practical observability and identifiability

Till now we have studied structural observability/identifiability. While structural identifiability is a property of the model structure, given a set of outputs, practical identifiability is related to the experimental data. In particular, it is a measure of the amount of information contained in these data.

A model can be structurally identifiable, but still be practically unidentifiable due to poor data quality, e.g., bad signal-to-noise ratio, errors in measurement or sparse sampling [61]. If structural identifiability is a prerequisite for

parameter identification, however this means that, if parameters are identifiable with ideal data (continuous, noise-free data) it does not guarantee that they will be practically identifiable with a finite number of noisy data points. Moreover estimation of parameters will use numerical algorithm and the the distance, for the problem considered, to the nearest ill-posed problem, [24, 38] i.e., conditioning of the problem will have an important role in the convergence of the algorithm.

In this section we will use sensitivity analysis and asymptotic statistical theory to study the practical identifiability. There are many good surveys and papers on this question [18, 7, 8, 22, 9].

We will begin by an intuitive introduction. As usual we consider the observed system

$$\begin{cases} \dot{x}(t) = f(x(t), \theta) \\ \dot{\theta} = 0 \\ y(t) = h(x(t), \theta). \end{cases} \quad (15)$$

with $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ and $\theta \in \mathbb{R}^p$.

We want to see what happens to the output for a small variation of the parameters (including the initial condition, considered as a peculiar case of parameter). By Linearization of the first order of the Taylor approximation, for a $\Delta\theta$ we can write

$$\Delta y(t, \theta) = \frac{\partial y}{\partial \theta}(t, \theta) \Delta\theta,$$

where $\frac{\partial y}{\partial \theta}(t, \theta)$ is the Jacobian of the observation with respect of the parameters (including the initial conditions). This Jacobian is a $m \times (2n + p)$ matrix and is called the sensitivity matrix

Reid introduces the notion of sensitivity identifiability if this relation can be solved uniquely in $\Delta\theta$. This linear problem is well known and if we denote by $A = \frac{\partial y}{\partial \theta}$ and if $\text{rank}(A) = 2n + p$ then the solution is given by means of the pseudo-inverse $A^+ = (A^T A)^{-1} A^T$

$$\Delta\theta = A^+ \Delta y = (A^T A)^{-1} A^T \Delta y.$$

It is also well known [34] that the sensitivity of this solution is ruled by the condition number $\kappa_2(A) = \frac{\sigma_{\max}}{\sigma_{\min}}$ with σ_{\max} and σ_{\min} respectively the greatest and smallest singular value of A (which are the corresponding eigenvalues of $A^T A$).

4.1 With Ordinary Least Squares

Now we assume that N observations has been obtained : Y_i $i = 1, \dots, N$ at times t_i . We assume that the obsevation is given by

$$Y_i = y(t_i, x_0, \theta) + \mathcal{E}_i,$$

with the error \mathcal{E}_i assumed to be a random variable satisfying the following assumptions

- The errors \mathcal{E}_i have mean zero $E[\mathcal{E}_i] = 0$;
- the errors have a common variance $\text{var}(\mathcal{E}_i)\sigma^2$;
- The error are i.i.d (independant and identically distributed)

The Fisher information matrix, for the preceding defined observations, is given by

$$FIM(\theta) = \sum_{i=1}^N \frac{\partial y}{\partial \theta}(t_i, \theta)^T \frac{\partial y}{\partial \theta}(t_i, \theta).$$

Solving the ordinary least square equations (OLS) gives an estimator $\hat{\theta}_{\text{OLS}}$ of the parameter θ (to simplify the notations, it is implied that we incorporate in θ the initial conditions) :

$$\hat{\theta}_{\text{OLS}} = \arg \min_{\theta} \sum_{i=1}^N [Y_i - y(t_i, x_0, \theta)]^2$$

Even though the error's distribution is not specified asymptotic statistical theory can be used to approximate the mean and variance of the random variable θ [67, 9] :

The bias adjusted approximation for σ^2 (with p parameters) is given by

$$\hat{\sigma}_{\text{OLS}}^2 = \frac{1}{N-p} \sum_{i=1}^N \left[Y_i - y(t_i, \hat{x}_0, \hat{\theta}_{\text{OLS}}) \right]^2.$$

This approximation of the variance is used to approximate the covariance matrix Σ

$$\Sigma = \hat{\sigma}_{\text{OLS}}^2 [\text{FIM}(\theta)]^{-1}.$$

The standard error for $\hat{\theta}_{\text{OLS}}$ can be approximated by taking the square roots of the diagonal elements of the approximation of the covariance matrix Σ

$$SE(\hat{\theta}_{\text{OLS}}(k)) = \sqrt{\Sigma_{kk}}$$

For example the 95% confidence interval for the k -th component of the parameter $\hat{\theta}_{\text{OLS}}(k)$ for p parameters, we use the Student's distribution with $N-p$ degrees of freedom

$$\hat{\theta}_{\text{OLS}}(k) - t_{0.025}^{N-p} \times SE(\hat{\theta}_{\text{OLS}}(k)) < \hat{\theta}_{\text{OLS}}(k) < \hat{\theta}_{\text{OLS}}(k) + t_{0.025}^{N-p} \times SE(\hat{\theta}_{\text{OLS}}(k))$$

From these formulas it appears that the conditioning of the Fisher Information Matrix plays an essential role. Huge confidence intervals give indications about the practicality of the identification.

The sensitivity matrix is obtained integrating an ODE.

$$\frac{\partial y}{\partial \theta}(t, x_0, \theta) = \frac{\partial}{\partial \theta} h(x(t, x_0, \theta)) = \frac{\partial h}{\partial x} \frac{\partial x}{\partial \theta}(t, x_0, \theta).$$

The matrix Jacobian $\frac{\partial h}{\partial x}$ is a $m \times n$ matrix. Now we have to evaluate the $n \times p$ matrix $\frac{\partial x}{\partial \theta}(t, x_0, \theta)$.

It is well known [36] that $\frac{\partial x}{\partial \theta}(t, x_0, \theta)$ is the solution of the linear equation

$$\frac{d}{dt} \left(\frac{\partial x}{\partial \theta}(t, x_0, \theta) \right) = \frac{\partial f}{\partial x}(x(t, x_0, \theta)) \left(\frac{\partial x}{\partial \theta}(t, x_0, \theta) \right) + \frac{\partial f}{\partial \theta}(x(t, x_0, \theta)).$$

In other words if we denote by $A(t) = \frac{\partial f}{\partial x}(x(t, x_0, \theta))$ the depending of time $n \times n$ matrix, by $B(t) = \frac{\partial f}{\partial \theta}(x(t, x_0, \theta))$ the depending of time $n \times p$ matrix,

then $z(t) = \left(\frac{\partial x}{\partial \theta}(t, x_0, \theta) \right)$ the $n \times p$ matrix is solution of following matrix linear equation

$$\dot{z}(t) = A(t) z(t) + B(t),$$

with the initial condition $z(0) = 0$.

For the dependency of the initial condition we have

$$\frac{\partial y}{\partial x_0}(t, x_0, \theta) = \frac{\partial}{\partial x} h(x(t, x_0, \theta)) = \frac{\partial h}{\partial x} \frac{\partial x}{\partial x_0}(t, x_0, \theta).$$

With the same reference [36] $\frac{\partial x}{\partial x_0}(t, x_0, \theta)$ is solution of the linear equation

$$\frac{d}{dt} \left(\frac{\partial x}{\partial x_0}(t, x_0, \theta) \right) = \frac{\partial f}{\partial x}(x(t, x_0, \theta)) \left(\frac{\partial x}{\partial x_0}(t, x_0, \theta) \right).$$

Denoting $w(t) = \frac{\partial x}{\partial x_0}(t, x_0, \theta)$, $w(t)$ is solution of the linear matrix equation

$$\dot{w}(t) = A(t) w(t),$$

with the initial condition $w(0) = Id_{\mathbb{R}^n}$.

To summarize we have to solve the following system in dimension $n^2 + np + n$

$$\left\{ \begin{array}{l} \dot{x}(t) = f(x(t), \theta) \\ \dot{z}(t) = A(t) z(t) + B(t) \\ \dot{w}(t) = A(t) w(t) \\ x(0) = x_0 \\ z(0) = 0_{n \times p} \\ w(0) = Id_{n \times n}. \end{array} \right. \quad (16)$$

With $A(t) = \frac{\partial f}{\partial x}(x(t, x_0, \theta))$ and $B(t) = \frac{\partial f}{\partial \theta}(x(t, x_0, \theta))$.

For big systems the computation of the different Jacobians can be prohibitive, in this case automatic differentiation software has to be used.

4.2 With Generalized Least Squares

With Ordinary Least Squares constant variance has been assumed which may be not appropriate for some data. A relative error, i.e., when the error is assumed to be proportional to the size of the measurement, is an assumption that might be reasonable when counting individuals in a population.

In this case we assume that the observation are [8, 20, 9, 15]:

$$Y_i = y(t_i, x_0, \theta) + y(t_i, x_0, \theta)^\rho \mathcal{E}_i.$$

The minimization is

$$\mathcal{L}(x_0, \theta) = \sum_{i=1}^N w_i [Y_i - y(t_i, x_0, \theta)]^2.$$

The values of the weight w_i depend on the value of the model and are not known. The process is carried with an iterated reweighted least squares

1. Estimate $\hat{\theta}_0$ and \hat{x}_0^0 with an OLS step : $w_1 = 1$
2. Set $\rho = 2$ and $w_i = 1 / [y(t_i, \hat{x}_0^0, \hat{\theta})]^{2\rho}$
3. Form $\mathcal{L}(x_0, \theta)$ with these w_j and estimate

$$(\hat{\theta}_1, \hat{x}_0^1) = \arg \min_{\theta, x_0} \mathcal{L}(x_0, \theta)$$

4. continue the procedure till the estimates $(\hat{\theta}_k, \hat{x}_0^k)$ are sufficiently close to each other, to obtain $\hat{\theta}_{GLS}$.

With $\hat{\theta}_{GLS}$, as in the OLS case, we can have an approximation of the correlation matrix, using the Fisher information matrix with weights

$$FIM((\hat{x}_0)_{GLS}, \hat{\theta}_{GLS}) = \sum_{i=1}^N \frac{1}{y(t_i, (\hat{x}_0)_{GLS}, \hat{\theta}_{GLS})^{2\rho}} \frac{\partial y}{\partial \theta}(t_i, (\hat{x}_0)_{GLS}, \hat{\theta}_{GLS})^T \frac{\partial y}{\partial \theta}(t_i, (\hat{x}_0)_{GLS}, \hat{\theta}_{GLS}).$$

Then we set

$$\Sigma_{GLS} = \hat{\sigma}_{GLS}^2 \left[FIM((\hat{x}_0)_{GLS}, \hat{\theta}_{GLS}) \right]^{-1},$$

with

$$\hat{\sigma}_{GLS}^2 = \frac{1}{N-p} \sum_{i=1}^N \frac{1}{y(t_i, (\hat{x}_0)_{GLS}, \hat{\theta}_{GLS})^{2\rho}} \left[Y_i - y(t_i, (\hat{x}_0)_{GLS}, \hat{\theta}_{GLS}) \right]^2.$$

The square roots of the diagonal elements of the approximation of the covariance matrix Σ_{GLS} give the standard errors.

4.3 Remarks

This parts use asymptotic statistical theory. The Kermack-McKenrik SIR model has been studied in a series of excellent papers [19, 18, 20, 15]. These papers approach the problem of observability/identifiability from the statistical point view : addressing parameter identifiability by exploiting properties of both the sensitivity matrix and uncertainty quantifications in the form of standard errors. Therefore the problem of the structural observability/identifiability is missed. For example in [20] the author identify $(\beta/N, \gamma, S, I)$ which we have proved to be identifiable. On the other hand in [15] the authors try to identify (S, I, N, β, γ) which is, as we have seen, hopeless. It is not surprising that the authors conclude : In this example, if we assume the initial conditions are known, our ability to estimate β and γ is good. Yet, once we have to estimate one or both initial conditions, our ability to estimate either β or γ worsens considerably. Given that in most situations initial conditions are not known exactly, parameter identifiability has the potential to be of widespread concern.

As already said a structural observability/identifiability must be done as a prerequisite. Then pursued by a sensitivity and uncertainty analysis. When doing this analysis the error structure of the data must be addressed. For example in [19] it is shown that using $\rho = 1$ in the GLS process gives better estimates for influenza data.

Another issue is when the output signal is not sufficiently informative (i.e., not persistently exciting [50]). For example when the data correspond to states near unobservability, e.g., near an equilibrium.

The problem of observability/ identifiability either theoretical (structural) or practical is not simple. More parameters are fitted, identifiability becomes a more serious problem.

References

- [1] D. AEYELS, *Generic observability of differentiable systems*, SIAM J. Control Optim., 19 (1981), pp. 595–603.
- [2] ———, *On the number of samples necessary to achieve observability*, Systems Control Lett., 1 (1981/82), pp. 92–94.
- [3] R. M. ANDERSON AND R. M. MAY, *Infectious Diseases of Humans. Dynamics and Control*, Oxford science publications, 1991.
- [4] M. ANGUELOVA, *Observability and identifiability of nonlinear systems with applications in biology*, PhD thesis, Chalmers University of Technology and Göteborg University, 2007.
- [5] S. AUDOLY, G. BELLU, L. D’ANGIO, M. P. SACCOMANI, AND C. COBELLI, *Global identifiability of nonlinear models of biological systems.*, IEEE Trans Biomed Eng, 48 (2001), pp. 55–65.
- [6] N. BACAER, *The model of Kermack and McKendrick for the plague epidemic in Bombay and the type reproduction number with seasonality*, J. Math. Biol., 64 (2012), pp. 403–422.
- [7] H. T. BANKS, A. CINTRÓN-ARIAS, AND F. KAPPEL, *Parameter selection methods in inverse problem formulation*, in Mathematical modeling and validation in physiology, vol. 2064 of Lecture Notes in Math., Springer, Heidelberg, 2013, pp. 43–73.
- [8] H. T. BANKS, M. DAVIDIAN, J. SAMUELS, AND K. SUTTON, *An inverse problem statistical methodology summary*, in Mathematical and Statistical Estimation Approaches in Epidemiology, G. Chowell, ed., Springer, 2009, pp. 249–302.
- [9] H. T. BANKS, S. HU, AND W. C. THOMPSON, *Modeling and inverse problems in the presence of uncertainty*, Monographs and Research Notes in Mathematics, CRC Press, Boca Raton, FL, 2014.

- [10] BELLMAN, R. AND ÅSTRÖM, K.J., *On structural identifiability*, Math. Biosci., 7 (1970), pp. 329–339.
- [11] G. BELLU, M. SACCOMANI, S. AUDOLY, AND L. D’ANGIO, *Daisy: a new software tool to test global identifiability of biological and physiological systems.*, Comput Methods Programs Biomed, 88 (2007), pp. 52–61.
- [12] F. BRAUER AND C. CASTILLO-CHÁVEZ, *Mathematical models in population biology and epidemiology*, vol. 40 of Texts in Applied Mathematics, Springer-Verlag, New York, 2001.
- [13] F. BRAUER, C. CASTILLO-CHAVEZ, AND Z. FENG, *Mathematical models in epidemiology*, vol. 69 of Texts in Applied Mathematics, Springer, New York, 2019. With a foreword by Simon Levin.
- [14] F. BRAUER, J. WU, AND P. VAN DEN DRIESSCHE, eds., *Mathematical Epidemiology*, no. 1945 in Lectures Notes in Math., Springer-Verlag, 2008.
- [15] A. CAPALDI, S. BEHREND, B. BERMAN, J. SMITH, J. WRIGHT, AND A. L. LLOYD, *Parameter estimation and uncertainty quantification for an epidemic model.*, Math Biosci Eng, 9 (2012), pp. 553–576.
- [16] O.-T. CHIS, J. R. BANGA, AND E. BALSACANTO, *Structural identifiability of systems biology models: a critical comparison of methods.*, PLoS One, 6 (2011), p. e27755.
- [17] CHIS, O. AND BANGA, J. R AND BALSACANTO, E., *GenSSI: a software toolbox for structural identifiability analysis of biological models.*, Bioinformatics, 27 (2011), pp. 2610–2611.
- [18] A. CINTRÓN-ARIAS, H. T. BANKS, A. CAPALDI, AND A. L. LLOYD, *A sensitivity matrix based methodology for inverse problem formulation*, J. Inverse Ill-Posed Probl., 17 (2009), pp. 545–564.
- [19] A. CINTRÓN-ARIAS, C. CASTILLO-CHÁVEZ, L. M. A. BETTENCOURT, A. LLOYD, AND H. T. BANKS, *The estimation of the effective reproductive number from disease outbreak data*, Tech. Rep. CRSC-TR08, Center for Research in Scientific Computation, 2008.

- [20] A. CINTRÓN-ARIAS, C. CASTILLO-CHÁVEZ, L. M. A. BETTENCOURT, A. L. LLOYD, AND H. T. BANKS, *The estimation of the effective reproductive number from disease outbreak data*, Math. Biosci. Eng., 6 (2009), pp. 261–282.
- [21] D. DALEY AND J. GANI, *Epidemic modelling : An introduction*, Cambridge University Press, 1999.
- [22] M. DAVIDIAN AND D. GILTINAN, *Nonlinear Models for Repeated Measurement Data.*, Chapman and Hall, Boca Raton, 1995.
- [23] M. C. DE JONG, O. DIEKMANN, AND H. HEESTERBEEK, *How does transmission of infection depend on population size ?*, in How does transmission of infection depend on population size? By: De Jong, Mart C. M.; Diekmann, Odo; Heesterbeek, Hans Conference: NATO Advanced Research Workshop on Epidemic Models Location: Cambridge, England, Publications of the Newton Institute, 1993, pp. 84–94.
- [24] J. W. DEMMEL, *On condition numbers and the distance to the nearest ill-posed problem*, Numer. Math., 51 (1987), pp. 251–289.
- [25] L. DENIS-VIDAL, G. JOLY-BLANCHARD, AND C. NOIRET, *Some effective approaches to check the identifiability of uncontrolled nonlinear systems*, Math. Comput. Simulation, 57 (2001), pp. 35–44.
- [26] S. DIOP AND M. FLIESS, *Nonlinear observability, identifiability, and persistent trajectories*, in proceedings 36th IEEE-CDC, 1991, pp. 714–719.
- [27] ———, *On nonlinear observability*, in Proceedings EEC91, vol. 1, Hermès, 1991, pp. 154–211.
- [28] M. C. EISENBERG, S. L. ROBERTSON, AND J. H. TIEN, *Identifiability and estimation of multiple transmission pathways in Cholera and waterborne disease.*, J Theor Biol, 324 (2013), pp. 84–102.
- [29] N. D. EVANS, L. J. WHITE, M. J. CHAPMAN, K. R. GODFREY, AND M. J. CHAPPELL, *The structural identifiability of the susceptible infected recovered model with seasonal forcing*, Math. Biosci., 194 (2005), pp. 175–197.

- [30] M. FLIESS, *Quelques définitions de la théorie des systèmes à la lumière des corps différentiels*, C. R. Acad. Sci. Paris Sér. I Math., 304 (1987), pp. 91–93.
- [31] ———, *Nonlinear control theory and differential algebra*, in Modelling and adaptive control (Sopron, 1986), vol. 105 of Lect. Notes Control Inf. Sci., Springer, Berlin, 1988, pp. 134–145.
- [32] ———, *Automatique et corps différentiels*, Forum Math., 1 (1989), pp. 227–238.
- [33] M. FLIESS AND T. GLAD, *Essays on Control*, no. 8 in Springer sciences, Springer, 1993, ch. An Algebraic Approach to Linear and Nonlinear Control, pp. 223–267.
- [34] G. H. GOLUB AND C. VAN LOAN, *Matrix Computations*, The John Hopkins University Press, 1989.
- [35] E. W. GRIFFITH AND K. S. P. KUMAR, *On the observability of nonlinear systems. I*, J. Math. Anal. Appl., 35 (1971), pp. 135–147.
- [36] P. HARTMAN, *Ordinary differential equations*, vol. 38 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Corrected reprint of the second (1982) edition [Birkhäuser, Boston, MA; MR0658490 (83e:34002)], With a foreword by Peter Bates.
- [37] R. HERMANN AND A. J. KRENER, *Nonlinear controllability and observability.*, IEEE Trans. Autom. Control, 22 (1977), pp. 728–740.
- [38] N. J. HIGHAM, *Matrix nearness problems and applications*, in Applications of matrix theory (Bradford, 1988), vol. 22 of Inst. Math. Appl. Conf. Ser. New Ser., Oxford Univ. Press, New York, 1989, pp. 1–27.
- [39] J. A. JACQUEZ AND P. GREIF, *Numerical parameter identifiability and estimability : Integrating identifiability, estimability and optimal sampling design.*, Math. Biosci., 77 (1985), pp. 201–277.
- [40] D. L. I. JANZEN, M. JIRSTRAND, M. J. CHAPPELL, AND N. D. EVANS, *Three novel approaches to structural identifiability analysis in mixed-effects models.*, Comput Methods Programs Biomed, (2016).

- [41] M. JIRSTRAND, *Algebraic Methods for Modeling and Design in Control*, PhD thesis, Linköping Studies in Science and Technology, Linköping University, 1996.
- [42] R. E. KALMAN, *Mathematical description of linear dynamical systems*, J. SIAM Control Ser. A, 1 (1963), pp. 152–192 (1963).
- [43] J. KARLSSON, M. ANGUELOVA, AND M. JIRSTRAND, *An efficient Method for Structural identifiability Analysis of Large Dynamic Systems*, in 16th IFAC Symposium on system identification, IFAC, 2012, pp. 941–946.
- [44] W. KERMAK AND A. MCKENDRICK, *A contribution to the mathematical theory of epidemics*, Proc. R. Soc., A115 (1927), pp. 700–721.
- [45] E. R. KOLCHIN, *Differential algebra and algebraic groups*, Academic Press, New York-London, 1973. Pure and Applied Mathematics, Vol. 54.
- [46] F.-X. LE DIMET AND O. TALAGRAND, *Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects*, Tellus, 38A (1986), pp. 97–110.
- [47] M. Y. LI, *An introduction to mathematical modeling of infectious diseases*, vol. 2 of Mathematics of Planet Earth, Springer, Cham, 2018.
- [48] J. LINTUSAARI, M. U. GUTMANN, S. KASKI, AND J. CORANDER, *On the identifiability of transmission dynamic models for infectious diseases.*, Genetics, 202 (2016), pp. 911–918.
- [49] L. LJUNG, *System identification : Theory for the user*, Prentice Hall, 1999.
- [50] L. LJUNG AND T. GLAD, *On global identifiability for arbitrary model parametrizations*, Automatica J. IFAC, 30 (1994), pp. 265–276.
- [51] P. MAGAL AND G. WEBB, *The parameter identification problem for SIR epidemic models: identifying unreported cases*, J. Math. Biol., 77 (2018), pp. 1629–1648.
- [52] M. MARTCHEVA, *An introduction to mathematical epidemiology*, vol. 61 of Texts in Applied Mathematics, Springer, New York, 2015.

- [53] H. MCCALLUM, N. BARLOW, AND J. HONE, *How should pathogen transmission be modelled?*, Trends Ecol Evol, 16 (2001), pp. 295–300.
- [54] N. MESHKAT, M. EISENBERG, AND J. J. DiSTEFANO, III, *An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner bases*, Math. Biosci., 222 (2009), pp. 61–72.
- [55] N. MESHKAT, Z. ROSEN, AND S. SULLIVANT, *Algebraic tools for the analysis of state space models*, in The 50th anniversary of Gröbner bases, vol. 77 of Adv. Stud. Pure Math., Math. Soc. Japan, Tokyo, 2018, pp. 171–205.
- [56] H. MIAO, X. XIA, A. S. PERELSON, AND H. WU, *On identifiability of nonlinear ODE models and applications in viral dynamics*, SIAM Rev., 53 (2011), pp. 3–39.
- [57] J. MURRAY, *Mathematical Biology I: An introduction*, vol. 17 of Interdisciplinary Applied Mathematics, Springer-Verlag, 2002.
- [58] V. K. NGUYEN, S. C. BINDER, A. BOIANELLI, M. MEYER-HERMANN, AND E. A. HERNANDEZ-VARGAS, *Ebola virus infection modeling and identifiability problems.*, Front Microbiol, 6 (2015), p. 257.
- [59] F. OLLIVIER, *Le problème de l’identifiabilité structurelle globale : étude théorique, méthodes effectives et bornes de complexité.*, PhD thesis, Ecole Polytechnique, 1990.
- [60] A. PERASSO, B. LAROCHE, Y. CHITOUR, AND S. TOUZEAU, *Identifiability analysis of an epidemiological model in a structured population*, J. Math. Anal. Appl., 374 (2011), pp. 154–165.
- [61] A. RAUE, C. KREUTZ, T. MAIWALD, J. BACHMANN, M. SCHILLING, U. KLINGMULLER, AND J. TIMMER, *Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood.*, Bioinformatics, 25 (2009), pp. 1923–1929.
- [62] J. F. RITT, *Differential algebra*, Dover Publications, Inc., New York, 1966.

- [63] W. C. RODA, M. B. VARUGHESE, D. HAN, AND M. Y. LI, *Why is it difficult to accurately predict the covid-19 epidemic?*, Infect Dis Model, 5 (2020), pp. 271–281.
- [64] M. SACCOMANI, S. AUDOLY, G. BELLU, AND L. D’ANGIÒ, *Parameter identifiability of nonlinear biological systems*, in Positive systems (Rome, 2003), vol. 294 of Lect. Notes Control Inf. Sci., Springer, Berlin, 2003, pp. 87–93.
- [65] M. P. SACCOMANI, *An effective automatic procedure for testing parameter identifiability of HIV/AIDS models*, Bull. Math. Biol., 73 (2011), pp. 1734–1753.
- [66] M. P. SACCOMANI, S. AUDOLY, AND L. D’ANGIÒ, *Parameter identifiability of nonlinear systems: the role of initial conditions*, Automatica J. IFAC, 39 (2003), pp. 619–632.
- [67] G. A. F. SEBER AND C. J. WILD, *Nonlinear regression*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1989.
- [68] A. SEDOGLAVIC, *A probabilistic algorithm to test local algebraic observability in polynomial time*, J. Symbolic Comput., 33 (2002), pp. 735–755.
- [69] E. D. SONTAG, *Mathematical control theory, deterministic finite dimensional systems*, no. 6 in Texts in Applied Mathematics, Springer-Verlag, 1990.
- [70] ———, *For differential equations with r parameters, $2r + 1$ experiments are enough for identification*, J. Nonlinear Sci., 12 (2002), pp. 553–583.
- [71] E. D. SONTAG AND Y. WANG, *I/o equations for nonlinear systems and observation spaces*, in Proc. IEEE Conf. Decision and Control, Brighton, UK., 1991, IEEE Publications, 1991, pages 720-725, 1991., 1991, pp. 720–725.
- [72] O. TALAGRAND, *On the mathematics of data assimilation*, Tellus, 33 (1981), pp. 321–339.
- [73] A. TARANTOLA, *Inverse problem theory and methods for model parameter estimation*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.

- [74] E. T. TUNALI AND T. J. TARN, *New results for identifiability of nonlinear systems*, IEEE Trans. Automat. Control, 32 (1987), pp. 146–154.
- [75] N. TUNCER, H. GULBUDAK, V. L. CANNATARO, AND M. MARTCHEVA, *Structural and practical identifiability issues of immuno-epidemiological vector-host models with application to Rift Valley Fever*, Bull. Math. Biol., 78 (2016), pp. 1796–1827.
- [76] N. TUNCER AND T. T. LE, *Structural and practical identifiability analysis of outbreak models*, Math. Biosci., 299 (2018), pp. 1–18.
- [77] N. TUNCER, M. MARCTHEVA, B. LABARRE, AND S. PAYOUTE, *Structural and practical identifiability analysis of Zika epidemiological models*, Bull. Math. Biol., 80 (2018), pp. 2209–2241.
- [78] A. F. VILLAVERDE, A. BARREIRO, AND A. PAPACHRISTODOULOU, *Structural identifiability of dynamic systems biology models.*, PLoS Comput Biol, 12 (2016), p. e1005153.
- [79] E. WALTER AND Y. LECOURTIER, *Global approaches to identifiability testing for linear and nonlinear state space models*, Math. Comput. Simulation, 24 (1982), pp. 472–482.
- [80] E. WALTER AND L. PRONZATO, *Identification de Modèles Paramétriques à partir de Données Expérimentales*, Masson, 1994.
- [81] E. WALTER AND L. PRONZATO, *Identification of parametric models*, Communications and Control Engineering Series, Springer-Verlag, Berlin; Masson, Paris, 1997. From experimental data, Translated from the 1994 French original and revised by the authors, with the help of John Norton.
- [82] H. WU, H. ZHU, H. MIAO, AND A. S. PERELSON, *Parameter identifiability and estimation of HIV/AIDS dynamic models*, Bull. Math. Biol., 70 (2008), pp. 785–799.
- [83] X. XIA AND C. H. MOOG, *Identifiability of nonlinear systems with application to HIV/AIDS models*, IEEE Trans. Automat. Control, 48 (2003), pp. 330–336.