



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. 2021. hal-02995562v2

HAL Id: hal-02995562

<https://hal.science/hal-02995562v2>

Preprint submitted on 6 Oct 2021 (v2), 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.

IDENTIFIABILITY AND OBSERVABILITY IN EPIDEMIOLOGICAL MODELS – A SURVEY *

F. HAMELIN[†], A. IGGIDR[‡], A. RAPAPORT[§], AND G. SALLET[¶]

Abstract. In this review, we recall the concepts of Identifiability and Observability of dynamical systems, and analyze them in the framework of Mathematical Epidemiology. We show that, even for simple and well known models of the literature, these properties are not always fulfilled. We then consider the problem of practical identifiability and observability, which are connected to sensitivity and numerical condition numbers. We also recall the concept of observers to reconstruct state variable of the model which are not observed, and show how it can be used with epidemiological models.

Key words. Dynamical systems, Identifiability, Observability, Mathematical Epidemiology, Sensitivity Analysis.

AMS subject classifications. 92D30, 93B30, 93B07, 93B53.

1. Introduction.

1.1. Prologue. Many papers in Mathematical 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 optimization algorithm with the use for example of a least square method or a maximum likelihood estimation. To validate the parameters estimation, one can use noisy synthetic simulated data obtained from the model for given values of the parameters, to check that the algorithm is able to reconstruct from the data the values of these parameters with accuracy.

One objective of this paper is to show that this procedure is not always safe 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, the SIR model of Kermack and Mckendrick [112]. This model got a renewed attention with the COVID-19 pandemic [72, 163]. Parameter identifiability analysis addresses the problem of which unknown parameters of an ODE model can uniquely be recovered from observed data. We will show that, even for very simple models, identifiability is far from being guaranteed.

The problem of identifiability for epidemiological models is relatively rarely addressed. For instance, a research in the Mathematical Reviews of the American Mathematical Society¹ with `epid*` AND `identifiability` gives only 4 papers, while

*Submitted to the editors October 6, 2021.

Funding: This work was not supported by any specific project.

[†]IGEPP, INRAE, Institut Agro, Univ Rennes, 35000, Rennes, France (frederic.hamelin@agrocampus-ouest.fr).

[‡]Université de Lorraine, CNRS, Inria, IECL, 57000 Metz, France (abderrahman.iggidr@inria.fr).

[§]MISTEA, Université Montpellier, INRAE, Institut Agro, 34060 Montpellier, France (alain.rapaport@inrae.fr).

[¶]Université de Lorraine, CNRS, IECL, 57000 Metz, France (gauthier.sallet@univ-lorraine.fr).

¹<https://mathscinet.ams.org/mathscinet>

`epidem*` AND `parameter` returns 68 publications. Only a small part of the later publications address the problem of identifiability. The following publications consider the problem of identifiability in epidemiological models: [10, 13, 24, 40, 41, 73, 77, 103, 119, 134, 140, 149, 167, 166, 187, 188, 189, 194, 204, 206]. However the majority of these papers is published elsewhere than in Biomathematics journals.

The question of observability, i.e. the ability to reconstruct state variables of the model from measurements, is often considered separately from the problem of identifiability. Either model parameters are known, or the identifiability analysis is performed prior to the study of observability. Indeed, the concepts of identifiability and observability are closely related, as we recall below. However, we shall show that for certain models, it is possible to reconstruct state variables with observers, while the model is not identifiable. In other situations, we show that considering jointly identifiability and observability with observers can be a way to solve the identifiability problem. This is another illustration of the interest of observers. This is why we shall dedicate a fair part of this review to the concept of observers and their practical constructions.

1.2. Definitions. The question of parameter identifiability originates from control theory and is related to observability and controllability [174]. The first appearance is in Kalman [109] and is now sixty years old. Identifiability is related to observability: the 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, which are also called the control, are considered as known. We will only consider systems without control, which is a peculiar case where the control is a singleton. When controls are known, with more information, observability/identifiability is sometimes easier. These problems have rarely been considered for uncontrolled systems whereas many methods have been developed for controlled systems. To be more precise, let us consider the following system in \mathbb{R}^n

$$(1.1) \quad \Sigma : \begin{cases} \dot{x}(t) &= f(x(t)), \\ x(0) &= x_0, \\ y(t) &= h(x(t)), \end{cases}$$

where we have denoted $\dot{x}(t) = \frac{dx}{dt}(t)$. The ordinary differential equation (ODE) $\dot{x} = f(x)$ is the dynamics 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 a 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 $t \geq 0$. This is often the case with epidemiological models for which the 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 any $t \geq 0$. 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 for an initial condition x_0 .

DEFINITION 1.1 (Observability). *The system (1.1) is observable if for two distinct initial conditions $x_1 \neq x_2$ 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 two distinct initial conditions, by considering only the observation.

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

$$(1.2) \quad \begin{cases} \dot{x}(t) &= f(x(t), \theta), \\ x(0) &= x_0, \\ y(t) &= h(x(t), \theta). \end{cases}$$

Identifiability is the ability to recover the unknown parameter from the observation. We denote by $x(t, x_0, \theta)$ the solution of (1.2) for an initial condition x_0 .

DEFINITION 1.2 (Identifiability). *System (1.2) is said to be identifiable if for any distinct parameters $\theta_1 \neq \theta_2$, there exists $t \geq 0$ such that*

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

There is an obvious similarity between observability and identifiability. Actually we will say that (1.2) is observable and identifiable if the augmented system

$$(1.3) \quad \begin{cases} \dot{x}(t) &= f(x(t), \theta), \\ \dot{\theta} &= 0, \\ y(t) &= h(x(t), \theta), \end{cases}$$

is observable.

Actually, for an epidemiological model it is unlikely to know the initial condition and it has long been recognized that initial conditions play a role in identifying the parameters [67, 121, 186, 206, 169].

What we have called identifiability is also known as *structural identifiability*. This expression has been coined by Bellman and K.J. Åström [23] 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 [103].

1.3. History. The observability concept has been introduced by Kalman [109] in the sixties for linear systems. For nonlinear systems, observability has been characterized circa the seventies [90, 94]. The definition is given by Hermann and Krener in the framework of differential geometry. Identifiability and structural identifiability has been introduced in compartmental analysis in 1970 by Bellman and Åström [23] in a paper that appeared in a bio-mathematics journal. The problem of identifiability is now addressed in text-books [120, 200, 198, 197]. Numerical identifiability of linear control system is implemented in softwares such as `Matlab` and `Scilab`.

Identifiability of nonlinear systems has been addressed in different context and the first systematic approach is by Tunali and Tarn in 1987 [186] in the differential geometry framework. The introduction of the concepts of differential algebra in control theory is due to Fliess around 1990 [67, 68, 79] followed by Glad [81, 121]. Identifiability is a general problem which has received different names depending on the community:

- observation, identification,
- data assimilation,
- inverse problem,
- parameters estimation.

“Data assimilation” is mainly used in en meteorology and oceanography [115, 181]. A direct (as opposed to inverse) 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 [183]. Finally, “parameters estimation” is used in the probability and statistics domains [4, 27, 116, 117, 144, 162].

1.4. Identifiability in mathematical epidemiology. Identifiability is well known in bio-mathematics from the seventies, as already mentioned with the paper of Bellman and Åström [23]. However, considering identifiability in mathematical epidemiology is relatively recent [187, 149, 206, 167, 134, 73, 77].

The first paper, to our knowledge, considering identifiability of an intra-host model of HIV is by Xia and Moog [206], and has been published in 2003 in a journal of automatic control.

1.5. The concept of observers. The construction of an *observer* is based on an estimation approach different from statistical methods: it consists of determining a dynamical system (called an “observer”) whose input is the vector $y(\cdot)$ of measures acquired over time, and whose state is an estimate $\hat{x}(t)$ of the (unknown) true state $x(t)$ of the system at time t .

An observer estimates $x(t)$ continuously over time and without anticipation, in the sense that the estimate $\hat{x}(t)$ is updated at each instant t through its dynamics as measurement $y(t)$ is available, without requiring the knowledge of any future measurement. This is why an observer is sometimes also called a “software sensor”. Since the estimate $\hat{x}(t)$ is given by the solution of a system of differential equations, the main idea behind an observer is the paradigm of *integrating instead of differentiating* the signal $y(\cdot)$. Note that although an observer is primarily devoted to state estimation, an observer can also aim to reconstructing simultaneously state and parameters, when some parameters are unknown (in this case a parameter vector p is simply considered a part of the system dynamics with $\dot{p} = 0$).

The most well-known observer is the so-called *Luenberger observer* [122] that is recalled in Section 4, and that has inspired most of the existing observers (several

ones are discussed in Section 4). However, observers are yet relatively unpopular in Mathematical Epidemiology, comparatively to other application domains (such as mechanics, aeronautics, automobile, *etc.*). The aim of the present review is also to promote the development and use of observers for epidemiological models.

Section 4 presents the theoretical background of observers construction and their convergence as estimators based on the model equations, independently of the quality of real data. In a complementary way, Section 5 discusses some implementation issues when observers are used with real world data that could be corrupted with noise.

2. Mathematical foundations.

2.1. Observability. We consider the following system

$$(2.1) \quad \Sigma \begin{cases} \dot{x} = f(x) \\ y = h(x) \\ x \in \mathbb{R}^n, y \in \mathbb{R}^m \end{cases},$$

where we assume, to avoid technical details, that f and h are \mathcal{C}^∞ functions. The function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is called a vector field.

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

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

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

2.1.1. Observability with Differential Geometry. The components of the observation map h are denoted by $h = (h_1, \dots, h_m)$. Each h_i is a \mathcal{C}^∞ function from the state space \mathbb{R}^n to \mathbb{R} .

DEFINITION 2.1 ([94]). *The observation space \mathcal{O} of (2.1) 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 .*

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

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

We have the following result

THEOREM 2.2. *For an analytic system (i.e., f and h are analytic 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

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

but, by induction we have the following relation

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

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

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

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 [94].

DEFINITION 2.3. *The system (2.1) 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 (2.1) 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 as follows

DEFINITION 2.4. [179] *Let*

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

where $d\psi(x)$ is the differential of ψ at x .

DEFINITION 2.5. *A system Σ is said to satisfy the observability rank condition in x if the dimension of $d\mathcal{O}$ in x satisfies*

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

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

THEOREM 2.6 (Hermann-Krener [94]). *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 non-singular Jacobian in x_0 .

As a consequence there exists an open set U containing x_0 where Φ is a bijection.

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 $\mathcal{L}_f^k h(x_0) = \mathcal{L}_f^k h(x_1)$ for any k therefore $\varphi_i(x_0) = \varphi_i(x_1)$ which is a contradiction.

PROPOSITION 2.7. *For an analytic system if the observability rank condition is satisfied everywhere the system is locally observable, hence observable.*

Proof. It is due to the fact that $h(x(t, x_0))$ is sum of its Taylor series. The rank condition implies, for the same reason as before, that the coefficients of the Taylor series separate points.

Example 2.8. We consider the SIR model of Kermack-McKendrick [112] for which the parameters β, γ are assumed to be known:

$$(2.2) \quad \begin{cases} \dot{S} &= -\beta \frac{S}{N} I \\ \dot{I} &= \beta \frac{S}{N} I - \gamma I \\ \dot{R} &= \gamma I \\ y &= \gamma I \end{cases},$$

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

With the gradients, the matrix

$$\begin{bmatrix} 0 & \gamma \beta \frac{I}{N} \\ \gamma & \gamma \left(\beta \frac{S}{N} - \gamma \right) \end{bmatrix}$$

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 Fliess proposed to use Differential Algebra for the analysis of nonlinear systems.

Intuitively, observability for system (2.1) is when the state x can be expressed as an algebraic expression of y and its time derivatives $y^{(k)}$ [68, 67].

DEFINITION 2.9. [68, 67] *A system is said to be algebraically observable if the states can be reconstructed by solving algebraic equations depending only of the observation and its time derivatives.*

Note that the systems under consideration are rational systems, i.e., the functions f and h are rational functions [67, 68]. A more precise definition can be given using the formalism of Differential Algebra. Differential Algebra can be considered as a generalization of the concepts of commutative algebra and algebraic geometry to differential equations. This theory, founded by Ritt, is an appropriate framework for defining algebraic observability. The interested reader can consult Fliess' publications for more details [78, 80].

Example 2.10. Consider the SIR model (2.2). This system evolves in the positively invariant set (which makes biological sense)

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

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

$$(2.3) \quad \Sigma \begin{cases} \dot{x} &= f(x) \\ y &= h(x) \end{cases},$$

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

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

It can be shown [10, 68, 171] that to obtain algebraic observability it is sufficient to consider the matrix

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

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

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

The number of Lie derivatives is bounded. Specifically, it is proved that it is sufficient compute no more than $n - 1$ derivatives [10, 68, 171].

THEOREM 2.11. [68, 179] *Consider (2.3) where f and h are rational functions. To check the observability rank condition, computing $(n - 1)$ Lie derivatives is sufficient.*

Proof. A detailed proof is given in the 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 $k + 1$ rational functions g_i , not all of them null, such that

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

Thanks to the definition of k , 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, $\mathcal{L}_f g_i$ is a rational function. Since $g_k \neq 0$ this prove that $d\mathcal{L}_f^{k+1} h$ is a linear combination of the previous Lie derivatives. This ends the proof. \square

The advantage of the Differential Algebra method promoted by Fliess and others is the possibility to be implemented inside a computer algebra software. Actually as soon as the system is in high dimension, the computations get rapidly involved. There is now software for local or global observability [171, 24, 40, 110, 194, 131, 97, 157, 15, 16].

Remark 2.12. For linear systems $\dot{x} = Ax$, $y = Cx$, all the definitions of observability are equivalent to having the *observability matrix*

$$(2.4) \quad O = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

full rank. This result can be extended to observation vector $y \in \mathbb{R}^p$ with $m > 1$ (see for instance [108]).

The observability analysis can also be a way to choose the right sensor, as illustrated on the following example.

Example 2.13. Consider a population model structured in five age classes, whose population sizes are

- x_1 for juveniles,
- x_2 for subadults capable of reproduction when adults,
- x_3 for subadults not capable of reproduction when adults,
- x_4 for adults capable of reproduction,
- x_5 for adults not capable of reproduction,

and the dynamics is

$$\begin{cases} \dot{x}_1 = -\alpha x_1 + \beta x_4 \\ \dot{x}_2 = \frac{\alpha}{2} x_1 - \alpha x_2 - m_1 x_2 \\ \dot{x}_3 = \frac{\alpha}{2} x_1 - \alpha x_3 - m_1 x_3 \\ \dot{x}_4 = \alpha x_2 - m_2 x_4 \\ \dot{x}_5 = \alpha x_3 - m_2 x_5 \end{cases}$$

(where α is an aging rate, m_1 , m_2 are mortality rates, and β is a fecundity rate). If only one sub-population x_i can be targeted for measurement, one can easily check that the only possibility for the system to be observable is to measure the variable $y = x_5$.

2.2. Identifiability. Since very often the initial conditions are not known, or partially known, we will consider in the following the problem of identifiability and observability, considering the augmented system (1.3). Note that identifiability-only problems are a special case in which $y = x$.

2.2.1. Characteristic set. We consider a polynomial system

$$(2.5) \quad \begin{cases} \dot{x}(t) = f(x(t), \theta) \\ x(0) = x_0 \\ y(t) = h(x(t), \theta) \end{cases},$$

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 belonging to $\mathbb{R}[x, y]$ (the set of polynomials in (x, y) with real coefficients). 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, an infinity of new equations. In other words the $n + m$ equations generate a differential ideal \mathcal{I} . Any x, y satisfying (2.5) 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 [113, 121, 159].

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 $\mathcal{C} \cap \mathbb{R}[\theta][x, y]$. This set is obtained by eliminating the state variables from the equations [63]. Actually, since we have no input, this characteristic set is the differential polynomial in y with the lower order [106]. This set is also called, in the literature, the input-output relation of the system [13, 24, 130, 105].

Making the polynomials monic in the input-output relations gives a set of coefficients $c_i(\theta)$ and a map $c : \mathbb{R}^p \rightarrow \mathbb{R}^\nu$ (for some ν) which is called the exhaustive summary of the model [142, 166, 169, 13].

The injectivity of c from the parameter space is only a necessary condition for identifiability [169]. Indeed, the input-output relations do not depend on the initial conditions and since the identifiability is dependent on the initial conditions it can happen that, with c injective, the system is not identifiable. Some authors [132, 187, 73] use the injectivity of c to ascertain the identifiability for almost all initial conditions. The following example shows that the injectivity of c is not sufficient for all initial conditions [169, Section 3.3].

Example 2.14 (input-output relation is not sufficient).

Consider the following compartmental system

$$(2.6) \quad \begin{cases} \dot{x}_1 &= -a_{21} x_1 + a_{12} x_2 \\ \dot{x}_2 &= -a_{12} x_2 \\ y &= x_1. \end{cases} .$$

We have

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

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 :

1. It is easy to see that

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

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

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

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

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

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

In the case of analytic systems an answer has been given by par D. Aeyels [3, 2] and E. Sontag [177]: 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. Let us stress that the whole state vector $x(\cdot)$ is assumed to be measured and therefore the initial state $x(0)$ is known.

2.3. Identifiability does not require observability. It can happen that a system is identifiable and however not observable.

Example 2.15. The following academic model is identifiable but not observable.

$$(2.7) \quad \begin{cases} \dot{x}_1 &= -\alpha(x_1 + x_2) \\ \dot{x}_2 &= \alpha(x_1 - x_2) \\ y &= \frac{1}{2}(x_1^2 + x_2^2) \end{cases}$$

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

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

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

2.4. Identifiability via change of coordinates. We show here that a change of variables can help to show the identifiability of a model. Let us consider a system in \mathbb{R}^n parameterized by $\theta \in \Theta \subset \mathbb{R}^p$

$$(2.8) \quad \begin{cases} \dot{x} &= f(x, \theta), & x(0) = x_0 \in X \\ y &= h(x) \end{cases}$$

where $X \subset \mathbb{R}^n$ is positively invariant for any $\theta \in \Theta$.

PROPOSITION 2.16. *Assume that the following properties hold.*

1. *The map f verifies*

$$(2.9) \quad \forall \theta_1, \theta_2 \in \Theta, \forall x \in X, \quad \theta_1 \neq \theta_2 \implies f(x, \theta_1) \neq f(x, \theta_2)$$

2. There exist smooth maps g, \tilde{g}, l such that

(a) for any solution $x(\cdot)$ of (2.8) in X , $w(t) := g(x(t), y(t)) \in W \subset \mathbb{R}^m$ verifies

$$(2.10) \quad \dot{w}(t) = l(w(t), y(t)), \quad t \geq 0$$

(b) for any $x \in X$, one has $w = g(x, h(x)) \iff x = \tilde{g}(w, h(x))$

Then the system (1.1) is identifiable over Θ for any initial condition in X .

Proof. Let $x_0 \in X$ and denote by $x_\theta(\cdot)$ the solution of (2.8) for the parameter $\theta \in \Theta$. Let θ_1, θ_2 be in Θ . If one has $h(x_{\theta_1}(t)) = h(x_{\theta_2}(t)) = y(t)$ for any $t \geq 0$, then the solution $w(\cdot)$ of $\dot{w} = l(w, y(t))$ for the initial condition $w(0) = g(x_0, y(0))$ verifies

$$\tilde{g}(w(t), y(t)) = x_{\theta_i}(t) \quad t \geq 0, \quad i = 1, 2,$$

that is $x_{\theta_1}(t) = x_{\theta_2}(t)$ for any $t > 0$. From hypothesis (2.9), we deduce that one has necessarily $\theta_1 = \theta_2$, which shows the identifiability of the system. \square

Let us illustrate this result on an intra-host model for malaria infection [25].

Exemple 2.17. The state vector is $x = (S, I_1, \dots, I_5, M)$ in \mathbb{R}_+^7 , where S is the concentration of uninfected erythrocytes in the blood, I_i are the concentrations of infected erythrocytes in different age classes, and M is the concentration of free merozoites. The dynamics is given by the following system

$$(2.11) \quad \begin{cases} \dot{S} = \Lambda - \mu_S S - \beta SM, \\ \dot{I}_1 = \beta SM - (\gamma_1 + \mu_1) I_1, \\ \dot{I}_2 = \gamma_1 I_1 - (\gamma_2 + \mu_2) I_2, \\ \vdots \\ \dot{I}_5 = \gamma_4 I_4 - (\gamma_5 + \mu_5) I_5, \\ \dot{M} = r \gamma_5 I_5 - \mu_M M - \beta SM, \end{cases}$$

where the different parameters are

Λ : recruitment of the healthy red blood cells (RBC).

β : rate of infection of RBC by merozoites.

μ_S : natural death rate of healthy cells.

μ_i : natural death rate of i -th stage of infected cells.

γ_i : transition rate from i -th stage to $(i+1)$ -th stage of infected cells.

r : number of merozoites released by the late stage of infected cells.

μ_M : natural death rate of merozoites.

The two first stages of infected erythrocytes (I_1 and I_2) correspond to the concentration of free circulating parasitized erythrocytes than can be observed (seen on peripheral blood smears) i.e. the quantity $I_1(t) + I_2(t)$ is measured at time t . The model (2.11) takes the form

$$(2.12) \quad \begin{cases} \dot{x} = Ax + \beta SM E + \Lambda e_1, \\ y = Cx \end{cases}$$

with

$$A = \begin{bmatrix} -\mu_S & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\gamma_1 - \mu_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma_1 & -\gamma_2 - \mu_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma_2 & -\gamma_3 - \mu_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma_3 & -\gamma_4 - \mu_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_4 & -\gamma_5 - \mu_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & r\gamma_5 & -\mu_M \end{bmatrix},$$

$$E = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \quad e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad C = [0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0]$$

Among parameters in (2.11), some of them (μ_i , γ_i , and r) are known or at least widely accepted by the community, but the infection rate β is unknown and cannot be estimated by biological considerations. Note that one has $ECE = E$. Therefore one can consider the variable $w = x - Ey = (I - EC)x$ whose dynamics is independent of the non-linear term βSM :

$$\dot{w} = \bar{A}w + \bar{A}Ey + \Lambda e_1$$

where we posit $\bar{A} = A - ECA$. The state x is then given by $x = w + Ey$. So we are exactly in the conditions of Proposition 2.16 for $\theta = \beta$ if one considers $X = (\mathbb{R}_+ \setminus \{0\})^7$. We immediately deduce that the parameter β is identifiable.

This example illustrates the interest of exploiting conjointly identifiability and observability to solve the identifiability problem.

3. Analysis of the classical SIR model of Kermack-McKendrick. The SIR model of Kermack and McKendrick [112] 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 [6, 30, 31, 32, 54, 118, 127, 137]. 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 the following ways:

- in [30, 31, 32] the model is fitted to the plague in Eyam (in the year 1666);
- in [54] the model is fitted to an influenza epidemic in England and Wales;
- in [118] a fitting is done with simulated noisy data;
- in [127, 30], in a chapter devoted to fitting epidemiological models to data, a SIR model is fitted to an influenza outbreak in an English boarding school.

More recently two publications [14, 124] revisit the fit of the Kermack-McKendrick SIR model to the plague in Bombay.

As already mentioned, before attempting to adjust parameters, an identifiability analysis should be performed.

3.1. The different forms of the SIR model. The original model [112] is

$$(3.1) \quad \begin{cases} \dot{S} &= -\tilde{\beta} S I \\ \dot{I} &= \tilde{\beta} S I - \gamma I \\ \dot{R} &= \gamma I \end{cases} ,$$

where S, I, R represent respectively the numbers of susceptible, infectious and removed individuals.

This model can also be found in a slightly different form

$$(3.2) \quad \begin{cases} \dot{S} &= -\beta \frac{S}{N} I \\ \dot{I} &= \beta \frac{S}{N} I - \gamma I \\ \dot{R} &= \gamma I \end{cases} ,$$

where $N = S + I + R$ is the total population. Obviously, one can pass from one model to the other though $\tilde{\beta} = \beta/N$. Both models are mathematically equivalent as long as N is a constant. However, we stress that identifying $\tilde{\beta}$ does not allow one to estimate the parameters β and N separately. For instance, estimating $\tilde{\beta}$ and γ only (without knowing N or β) does not allow one to estimate the basic reproduction number:

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

3.2. Observability-identifiability of the SIR model. Quite surprisingly, the observability and identifiability of the original Kermack-Mckendrick SIR model has not been studied much, although this model is commonly used to model epidemics.

Interestingly, the observability and identifiability of the SIR model with births and deaths, constant population, and an observation $y = kI$, has been first studied in 2005 [77]:

$$(3.3) \quad \begin{cases} \dot{S} &= \mu N - \beta \frac{S}{N} I - \mu S \\ \dot{I} &= \beta \frac{S}{N} I - (\gamma + \mu) I \\ \dot{R} &= \gamma I - \mu R \end{cases} ,$$

where μ is the renewal rate of the population. The article [77] concludes that the system is neither observable nor identifiable.

In [188] the identifiability of (3.2) is addressed assuming (i) that the initial conditions (and therefore $N = S(0) + I(0) + R(0)$) are known, (ii) observing $y = kI$ with $k = 1$, and using only the input-output relation to conclude. Under assumptions (i) and (ii), the identifiability is quite immediate, as we shall see, but of limited interest.

Consider the SIR model

$$(3.4) \quad \begin{cases} \dot{S} &= -\beta \frac{S}{N} I \\ \dot{I} &= \beta \frac{S}{N} I - \gamma I \\ y &= k I \end{cases} .$$

The last equation has been omitted since $R = N - S - I$. The observation is $y = k I$, in other words only a fraction of the infectious individuals are observed. This situation is used for example in [124, 163].

THEOREM 3.1. *The variables and parameters to be estimated are S, I, N, β, γ , and k .*

System (3.4) 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.

Remark 3.2. One could believe that if $k = \gamma$, with N unknown, then (3.1) is observable. This is wrong. Certainly S, I are observable, but not R . Therefore $N = S + I + R$ is inaccessible. As a consequence, $\mathcal{R}_0 = \beta N / \gamma$ is not identifiable.

Theorem 3.1 can be obtained from [77] by setting $\mu = 0$. However we provide a short and elementary proof.

Proof. We will show that the parameters and the state can be expressed with the observation and their derivatives [68, 67]. Otherwise the identifiability will not be obtained. We consider the system on

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

This open set \mathcal{U} is positively invariant. At 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 $I \neq 0$)

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

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

Therefore the only information 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 $kS, kI, \gamma, \frac{\beta}{kN}$ are identifiable. \square

3.2.1. Using input-output relations. We proceed by elimination:

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

Equivalently, one has

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

We have already seen, with the notation of the previous proof, that

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

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

3.2.2. Using ORC Herman-Krener condition. Consider the application

$$\Phi : (S, I, \beta, \gamma, k) \longrightarrow (y, y^{(1)}, y^{(2)}, y^{(3)}, y^{(4)})$$

and denote by

$$Jac \Phi = \frac{\partial \Phi}{\partial (S, I, \beta, \gamma, k)}$$

the Jacobian of Φ . Then $\det(Jac \Phi) = 0$ proves that the system is neither identifiable nor observable.

On the other hand with

$$\Psi : (S, I, \beta, \gamma) \longrightarrow (y, y^{(1)}, y^{(2)}, y^{(3)})$$

and denoting by

$$Jac \Psi = \frac{\partial \Psi}{\partial (S, I, \beta, \gamma)}$$

the Jacobian of Ψ , 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 observation are the cumulative numbers of infectious individuals. We study how this change the observability and identifiability of the SIR model.

We consider the system where the observation is given by

$$(3.5) \quad y(t) = k \int_0^t \beta \frac{S(\tau) I(\tau)}{N} d\tau.$$

This problem has been addressed for the SIR model with demography for constant population in [77]. Identifiability with known initial conditions for (3.4) is also considered in [188] using input-output relations.

THEOREM 3.3. *The system (3.4) with cumulative incidence observation is neither observable, nor identifiable.*

The parameters $kS, kI, \gamma, \frac{\beta}{kN}$ are identifiable.

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

Proof. A straightforward computation gives:

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

Finally

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

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

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

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.

We have

$$kS - kI = \frac{\gamma + h_2}{\frac{\beta}{kN}}, \quad kS - kI = \frac{\dot{y}}{\frac{\beta}{kN}}.$$

□

which proves our claim.

4. Observers synthesis. So far we have studied observability as a property such that a measured “signal” $y(\cdot)$ allows to reconstruct the state of the system univoquely, by studying the information provided by the successive derivatives of $y(\cdot)$. Formally, when the system is observable and we know (perfectly) enough derivatives of $y(\cdot)$ at a given time t , we just have to invert the map $\phi \mapsto (h(x), \mathcal{L}_f h(x), \dots)$ at $(y(t), \dot{y}(t), \ddot{y}(t), \dots)$ to reconstruct the state variable $x(t)$. In practice, it is known that numerically calculating derivatives on the raw signal data is imprecise and sensitive to measurement noise, especially if several successive derivatives have to be determined. It is generally preferable to use a “filter” to smooth the data, for example using polynomial functions or splines that approach with more regularity the measurements obtained over time, on which the derivative calculations can be performed before the inversion operation.

In this section, we study the construction of observers and their theoretical convergence, without considering their practical performances in presence of measurements noise. This point will be addressed in the more practical Section 5.

Observers for dynamics $\dot{x} = f(x)$ with observation $y = h(x)$ of the form

$$\dot{\hat{x}} = f(\hat{x}) + G((h(\hat{x}) - y(t)))$$

where G is a constant matrix, are called *Luenberger observers* [122]. Note that this construction consists in a copy of the original dynamics f plus a correcting term

which depends on an *innovation* term, that is the difference between the expected output $\hat{y} = h(\hat{x}(t))$ if the true state was $\hat{x}(t)$ and the effective measured output $y(t)$. Therefore, if $\hat{x}(t)$ and $x(t)$ are equal at a certain time t , it will remain identical at any future time. The main point is that the matrix G has to be chosen such that the estimation error $\hat{x}(t) - x(t)$ does converge to 0. When f and h are linear and the system is observable, the theory of linear automatic control teaches that there always exists G such that the convergence speed of the estimator can be chosen arbitrarily fast (see e.g. [8]). Obviously epidemiological models are rarely linear. However, looking for an Luenberger observer is often a first trial before considering more sophisticated estimators. Indeed, we shall see that for certain nonlinear dynamics, such observers do the job and in other cases, observers can be inspired from this form.

Let us begin by some simple cases of observers for particular dynamics, and then we shall present a more general framework.

4.1. Observers with linear error dynamics. Consider systems of the form

$$(4.1) \quad \begin{cases} \dot{x} &= Ax + \Phi(t, y) \\ y &= Cx \end{cases}$$

where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$. Then, observers of the form

$$(4.2) \quad \dot{\hat{x}} = A\hat{x} + \Phi(t, y(t)) + G(C\hat{x} - y(t))$$

are such that the error vector $e(t) = \hat{x}(t) - x(t)$ is solution of the linear dynamics

$$(4.3) \quad \dot{e} = (A + GC)e$$

The choice of the *gains* vector G providing a convergence of $e(t)$ to 0 comes directly from the *poles placement* technique of the theory of linear systems that we recall below.

The following lemma is well-known and often used in automatic control.

LEMMA 4.1. *Let A be a square matrix of size n and C be a line vector of length n . If the observability matrix \mathcal{O} defined in (2.4) is of full rank, then for any set $\Lambda = \{\lambda_1, \dots, \lambda_n\}$ of n real or complex two-by-two conjugate numbers, there is a G vector of size n such that*

$$Sp(A + GC) = \Lambda.$$

Specifically, if

$$\pi_A(\xi) = \xi^n + a_1\xi^{n-1} + \dots + a_{n-1}\xi + a_n$$

is the characteristic polynomial of A , then one has

$$G = P \left[a_{n+1-i} + (-1)^{n-i} \sigma_{n+1-i}(\Lambda) \right]_{i=1}^n$$

where

$$P = O^{-1} \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{bmatrix} [I \ A \ \dots \ A^{n-1}]$$

and the σ_k designate the symmetric functions of the roots

$$(4.4) \quad \sigma_k(\Lambda) = \sum_{1 \leq i_1 \leq \dots \leq i_k \leq n} \lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_k}$$

A proof of Lemma 4.1 is given in Appendix A.

Remark. This result can be generalized to vectorial observations, i.e. for matrices C with $m > 1$ rows and n columns.

Finally, by choosing numbers λ_i with negative real parts, one can make the convergence of the error given by the exponentially decreasing dynamics (4.3) as fast as desired. The observer (4.2) is thus adjustable. It should be noted that when the difference $\hat{z}_1(t) - y(t)$ (usually called "innovation") becomes and remains close to 0, the trajectories of the observer follow those of the system: we can then consider that the observer has practically converged. The innovation is thus very useful in practice because it provides information on the current stage of convergence of the estimate.

We illustrate this technique on a population model with age classes.

Example 4.2. Let us consider a population structured in three stages: young, subadult and adult, of stocks x_1 , x_2 , x_3 respectively. It is assumed that only adults x_3 can reproduce, giving birth to young x_1 :

$$(4.5) \quad \begin{cases} \dot{x}_1 &= -a_1x_1 - m_1x_1 + r(t, x_3) \\ \dot{x}_2 &= a_1x_1 - a_2x_2 - m_2x_2 \\ \dot{x}_3 &= a_2x_2 - m_3x_3 \\ y &= x_3 \end{cases} .$$

The coefficients a_i are the transition rates between age classes, m_i are the mortality rates of each class, and $r(\cdot)$ is the reproductive function (usually non-linear and seasonally dependent), for example

$$r(t, x_3) = \frac{\bar{r}(t)x_3}{k + x_3}, \quad \bar{r}(t) \in [\bar{r}_{min}, \bar{r}_{max}] .$$

Here, it is assumed that only the size of the adult class is measured over time. The aim is to estimate the stocks of larvae and subadults over time. The model (4.5) is of the form (4.1), where we posed

$$A = \begin{bmatrix} -(a_1 + m_1) & 0 & 0 \\ a_1 & -(a_2 + m_2) & 0 \\ 0 & a_2 & -m_3 + 0 \end{bmatrix}, \quad C = [0 \ 0 \ 0 \ 1]$$

and

$$\Phi(t, x_3) = \begin{bmatrix} 0 \\ r(t, x_3) \end{bmatrix}$$

One can check that the observability matrix \mathcal{O} is full rank, or alternatively directly check that the system is observable. Indeed, we obtain by using the expression \dot{x}_3 :

$$x_2 = \frac{\dot{y} + m_3y}{a_2},$$

then with the expression \dot{x}_2 :

$$x_1 = \frac{\dot{x}_2 + (a_2 + m_2)x_2}{m_1} = \frac{\ddot{y} + (a_2 + m_2)m_3y}{a_2m_2} .$$

Therefore, the following system

$$(4.6) \quad \begin{cases} \dot{\hat{x}}_1 &= -a_1\hat{x}_1 - m_1\hat{x}_1 + r(t, y(t)) + G_1(\hat{x}_3 - y(t)) \\ \dot{\hat{x}}_2 &= a_1\hat{x}_1 - a_2\hat{x}_2 - m_2\hat{x}_2 + G_2(\hat{x}_3 - y(t)) \\ \dot{\hat{x}}_3 &= a_2\hat{x}_2 - m_3\hat{x}_3 + G_3(\hat{x}_3 - y(t)) \end{cases}$$

with well chosen G_1, G_2, G_3 numbers is an observer for the dynamics (4.5), with an exponential convergence.

It may also happen that the estimation error of an observer is only partially assignable, as we shall see in the next example.

4.2. Observers via change of coordinates. Consider, as in Section 2.4, systems

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

for which there exists a change of coordinates $x \mapsto w = g(x, h(x))$ such that one has

$$\dot{w}(t) = l(w(t), h(x(t))), \quad t \geq 0$$

for any solution $x(\cdot)$, with the properties

1. $w = g(x, h(x)) \iff x = \tilde{g}(w, h(x))$
2. $w = g(x, h(x)) \implies h(x) = k(w)$

where g, \tilde{g}, l and k are smooth maps.

Then, one can look for an observer $\hat{w}(\cdot)$ of the system

$$\begin{cases} \dot{\hat{w}} &= l(w, h(x)) \\ y &= k(w) \end{cases}$$

and take, as an estimator of $x(\cdot)$

$$(4.7) \quad \hat{x}(t) = \tilde{g}(\hat{w}(t), y(t)).$$

There is an advantage of considering such a change of coordinates when the maps g, \tilde{g}, l and k are independent of a parameter θ present in the expression of f , as in Proposition 2.16 of Section 2.4. However, the estimator (4.7) does not filter the measurement $y(\cdot)$ and might be sensitive to noise.

Let us illustrate this approach on the malaria model (Example 2.17).

Example 4.3. Consider the model (2.11) of Example 2.17. With the variable $w = x - Ey = (I - EC)x$, the dynamics (2.12) is independent of the unknown non-linear term βSM :

$$\begin{cases} \dot{w} &= \bar{A}w + \bar{A}Ey + \Lambda e_1 \\ y &= Cx = Cw \end{cases}$$

and one can consider the following observer for system (2.12)

$$(4.8) \quad \begin{cases} \dot{\hat{w}} &= (\bar{A} - LC)\hat{w} + (L + (\bar{A} - LC)E)y(t) + \Lambda e_1, \\ \hat{x}(t) &= \hat{w}(t) + Ey(t). \end{cases}$$

where L is a vector in \mathbb{R}^7 to be chosen. The dynamics of the error $e(t) = \hat{x}(t) - x(t)$ is given by

$$\dot{e} = (\bar{A} - LC)e$$

Note that one has $C\bar{A} = 0$. Therefore the rank of the observability matrix of the pair (\bar{A}, C) is equal to one, and 0 is an eigenvalue of \bar{A} . The choice of L allows then to assign only one eigenvalue of $\bar{A} - LC$, equal to $-(L_2 + L_3)$, the other eigenvalues remaining negative. Therefore (4.8) is an observer for system (2.12) with exponential convergence, that does not use the unknown parameter β .

Remark 4.4. Differently to observers of Section 4.1, one cannot expect a convergence speed of the observer (4.8) faster than the dynamics (2.11), because the error dynamics is not completely assignable. However, the convergence is exponential. This is illustrated with numerical simulations in Section 5.2.2.

4.3. Reduced-order observers. A typical situation is when one can operate a state decomposition as follows

1. decompose the state vector x (may be at the price of a change of variables) as

$$x = \begin{bmatrix} y \\ x_u \end{bmatrix}$$

where x_u represent the *unmeasured* variables,

2. look for an auxiliary variable (that we called z) whose dynamics is independent of x_u

$$\dot{z} = g(z, y)$$

and asymptotically stable, and such that x_u can be globally expressed as

$$x_u = l(z, y)$$

where l is a smooth map (say C^1).

Then, the dynamics

$$\begin{cases} \dot{\hat{z}} &= g(\hat{z}, y(t)) \\ \hat{x}_u &= l(\hat{z}, y(t)) \end{cases}$$

is an *asymptotic observer*, whose error convergence $\hat{x}_u - x_u$ is simply provided by the asymptotic convergence of \hat{z}_z to 0, whatever is the initial condition $\hat{z}(0)$. When the convergence speed of an estimator cannot be adjusted, it is usually called an *asymptotic observer*, differently to the previous section for which the error convergence can be made arbitrarily fast. Note that differently to the previous section, these observers have no tuning parameters and are not driven by innovation terms. These estimators are *reduced-order* observers when the variable z is of lower dimension than x . An interest for such observers is that it can possess robustness features when the maps g and l are independent of some terms or parameters of the dynamics $\dot{x} = f(x)$. Let us illustrate this feature on the Kermack-McKendrick model with fluctuating rates.

Example 4.5. We consider the SIR model with birth and death terms

$$(4.9) \quad \begin{cases} \dot{S} &= -\beta(t)SI + \nu N - \mu S \\ \dot{I} &= \beta(t)SI - \rho(t)I - \mu I \\ \dot{R} &= \rho(t)I - \mu R \end{cases}$$

where parameters β and ρ fluctuate unpredictably over time. We assume, for simplicity, that the birth rate ν is equal to the death rate μ , so that the total population remains constant of size $N = S + I + R$ (assumed to be known). Let us suppose that

the size of the infected population is monitored over time as well as the number of new cured individuals, which amounts considering that the observation vector at time t is

$$y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} I(t) \\ \rho(t)I(t) \end{bmatrix}.$$

Stocks of classes S and R are not initially known. Then, the system

$$(4.10) \quad \begin{cases} \dot{Z} &= \nu N - y_2(t) - \mu Z \\ \dot{\hat{S}} &= Z - y_1(t) \\ \dot{\hat{R}} &= N - Z \end{cases}$$

is an observer allowing to estimate S and R without knowing $\beta(\cdot)$ and $\rho(\cdot)$. Indeed, the dynamics of the estimators is

$$\begin{aligned} \frac{d}{dt}(\hat{S} - S) &= \dot{\hat{S}} - \dot{S} = -\mu(Z - S - I) = -\mu(\hat{S} - S) \\ \frac{d}{dt}(\hat{R} - R) &= -\dot{Z} - \dot{R} = -\nu N + \mu(Z + R) = -\mu(\hat{R} - R) \end{aligned}$$

which ensures the convergence of the \hat{S} and \hat{R} estimates. Note that the internal dynamics of the observer is here of smaller dimension than the system, and that the estimate of the unmeasured state variable S is a function of the internal state Z of the observer and the observation y_1 . The speed of convergence of this observer is not adjustable, but it has the advantage of being perfectly robust to any (unknown) variations of the terms $\beta(\cdot)$ and $\rho(\cdot)$. This is illustrated with numerical simulations in Section 5.2.3.

4.4. A general non-linear framework. In a general way, we will retain that an observer for a system

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

is an input-output system (input: y , output: \hat{x}) of the form

$$\begin{cases} \dot{\xi} &= g(t, \xi, y(t)) \\ \hat{x} &= l(t, \xi, y(t)) \end{cases}$$

such as the coupled system

$$\begin{cases} \dot{x} &= f(t, x) \\ \dot{\xi} &= g(t, \xi, h(t, x)) \end{cases}$$

verifies the property

$$\lim_{t \rightarrow +\infty} l(t, \xi(t), h(t, x)) - x(t) = 0$$

for any initial condition $x(0)$, $\xi(0)$.

In the two previous examples, the dynamics of the estimation error was linear. For an observable non-linear system, the existence of an observer whose estimation error is linear is not guaranteed. This is a difficult problem. However, one can consider

the (nonlinear) *observability canonical form* [85] (given here for a scalar output)

$$(4.11) \quad \dot{z} = F(z) := \underbrace{\begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ & & \ddots & \ddots & \\ & & & \ddots & \\ & & & & 0 & 1 \\ & & & & & & 0 \end{bmatrix}}_A z + \underbrace{\psi(z)}_B$$

$$(4.12) \quad y = \underbrace{[1 \ 0 \ \cdots \ \cdots \ 0]}_C z$$

where the map ψ is Lipschitz on \mathbb{R}^n . Then, one can show that there exists an observer of the Luenberger form

$$(4.13) \quad \dot{\hat{z}} = F(\hat{z}) + G(C\hat{z} - y(t))$$

with exponential convergence when G is a well-chosen gains vector. When an observable system

$$\begin{cases} \dot{x} = f(x), & x \in \mathbb{R}^n \\ y = h(x), & y \in \mathbb{R} \end{cases}$$

is not in normal form, but the application

$$\phi(x) = \begin{bmatrix} h(x) \\ \mathcal{L}_f h(x) \\ \vdots \\ \mathcal{L}_f^{n-1} h(x) \end{bmatrix}$$

is a diffeomorphism from \mathbb{R}^n into \mathbb{R}^n and the map

$$\psi(z) := \mathcal{L}_f^m h \circ \phi^{-1}(z)$$

is Lipschitz on \mathbb{R}^n , then the observer (4.13) can be written in the x coordinates as follows

$$\dot{\hat{x}} = f(\hat{x}) + [J\phi(\hat{x})]^{-1} G(h(\hat{x}) - y(t))$$

where $J\phi(x)$ denotes the Jacobian matrix of ϕ at x . The observer preserves the Luenberger structure but with variable gains.

Let us first note that the pair (A, C) as defined in (4.11)-(4.12) is observable. Indeed, we have $O = Id$. Thus, according to the Lemma 4.1, one can freely assign the spectrum of $A + GC$ by choosing the vector G . We show now how to choose the eigenvalues of $A + GC$ to ensure the convergence of the non-linear observer 4.13. To do this, we begin by giving some properties of the Vandermonde matrices

$$V_{\lambda_1, \dots, \lambda_n} := \begin{bmatrix} \lambda_1^{n-1} & \lambda_1^{n-2} & \cdots & \lambda_1 & 1 \\ \lambda_2^{n-1} & \lambda_2^{n-2} & \cdots & \lambda_2 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \lambda_n^{n-1} & \lambda_n^{n-2} & \cdots & \lambda_n & 1 \end{bmatrix}$$

related to the normal form.

LEMMA 4.6. *Let $\Lambda = \{\lambda_1, \dots, \lambda_n\}$ be a set of n distinct real numbers and G a vector such that*

$$Sp(A + GC) = \Lambda := \{\lambda_1, \dots, \lambda_n\}$$

Then

$$V_{\lambda_1, \dots, \lambda_n}(A + GC)V_{\lambda_1, \dots, \lambda_n}^{-1} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

Moreover, for any $C > 0$ and $\theta > 0$, there exist $\lambda_n < \lambda_{n-1} < \dots < \lambda_1 < 0$ such that

$$\lambda_1 + C\|V_{\lambda_1(\theta), \dots, \lambda_n(\theta)}^{-1}\|_\infty = -\theta$$

The proof of Lemma 4.6 is given in Appendix B.

We are now ready to show the convergence of the observer (4.13) in coordinates z , for a gains vector G such that $A + GC$ has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$ of negative real values. Denote the error $e = \hat{z} - z$. We have

$$\dot{e} = (A + GC)e + B(\psi(\hat{z}) - \psi(z))$$

Let $\xi = Ve$ where V designates the Vandermonde matrix $V_{\lambda_1, \dots, \lambda_n}$. Thanks to Lemma 4.6, we obtain

$$\dot{\xi} = \Delta\xi + VB(\psi(\hat{z}) - \psi(z))$$

where Δ is the diagonal matrix $diag(\lambda_1, \dots, \lambda_n)$. By multiplying on the left by ξ^\top , one obtains

$$\begin{aligned} \xi^\top \dot{\xi} &= \xi^\top \Delta\xi + \xi^\top VB(\psi(\hat{z}) - \psi(z)) \\ &\leq \lambda_1 \|\xi\|^2 + \|\xi\| \sqrt{n} |\psi(\hat{z}) - \psi(z)| \\ &\leq \lambda_1 \|\xi\|^2 + \|\xi\| \sqrt{n} L \|e\| \\ &\leq (\lambda_1 + \sqrt{n} L \|V^{-1}\|_\infty) \|\xi\|^2 \end{aligned}$$

where L is the Lipschitz constant of ψ . Thus the norm of ξ verifies

$$\|\xi(t)\| \leq \|\xi(0)\| + \int_0^t (\lambda_1 + \sqrt{n} L \|V^{-1}\|_\infty) \|\xi(\tau)\| d\tau$$

and by Gronwall's Lemma, we obtain

$$\|\xi(t)\| \leq \|\xi(0)\| e^{(\lambda_1 + \sqrt{n} L \|V^{-1}\|_\infty)t}$$

Finally, for any $\theta > 0$, Lemma 4.6 gives the existence of numbers $\lambda_n < \lambda_{n-1} < \dots < \lambda_1 < 0$ such that $\|\xi(t)\| \leq \|\xi(0)\| e^{-\theta t}$ for any $t > 0$, which guarantees the exponential convergence of the error e to 0.

The observer (4.13) with the gains vector G_θ is called *high gain observer* [85], because the value of θ must be "sufficiently" large, and its successive powers might take large values.

Remark 4.7. In practice, the map ψ is not necessarily globally Lipschitz on \mathbb{R}^n . Nevertheless, if there exists a compact sub-set K of \mathbb{R}^n that is forwardly invariant by the dynamics (4.11), one can consider an extension of ψ outside K that is globally Lipschitz on \mathbb{R}^n and define then the observer on whole \mathbb{R}^n (see [152]).

Let us now illustrate this construction on the Kermack-McKendrick model model.

Exemple 4.8. We consider the classical SIR model

$$(4.14) \quad \begin{cases} \dot{S} &= -\beta SI \\ \dot{I} &= \beta SI - \rho I \\ \dot{R} &= \rho I \end{cases}$$

where the parameters β and ρ are known, and suppose that the only observation is the cumulative number of recovered individuals since a time t_0

$$y(t) = \int_{t_0}^t \rho I(\tau) d\tau.$$

It is also assumed that the size of the total population $N = S + I + R$ is known. To put the system in canonical form, we write

$$\begin{aligned} z_1 &= y \\ z_2 &= \dot{z}_1 = \rho I \\ z_3 &= \dot{z}_2 = (\beta S - \rho) z_2 \end{aligned}$$

and

$$\dot{z}_3 = (\beta S - \rho) z_3 - \beta(\beta SI) z_2 = \psi(z) := \frac{z_3^2}{z_2} - \frac{\beta}{\rho} \left(\frac{z_3}{z_2} + \rho \right) z_2^2.$$

We can then reconstruct the I , S and R stocks from the variables z as follows

$$\begin{aligned} S &= \frac{1}{\beta} \left(\frac{z_3}{z_2} + \rho \right) \\ I &= \frac{z_2}{\rho} \\ R &= N - \frac{1}{\beta} \left(\frac{z_3}{z_2} + \rho \right) - \frac{z_2}{\rho} \end{aligned}$$

Note that ψ is not globally Lipschitz on \mathbb{R}^3 , and has a singularity at $z_2 = 0$. Nevertheless, we notice that the term z_3/z_2 can be framed as follows

$$\frac{z_3}{z_2} = \beta S - \rho \in [-\rho, \beta - \rho]$$

and that one has

$$\dot{z}_3 \in [-\rho^3 N - \rho\beta^2 N^3, \rho(\beta - \rho)N].$$

We can therefore consider the expression

$$\tilde{\psi}(z) = \text{sat}_{[-\rho^3 N - \rho\beta^2 N^3, \rho(\beta - \rho)N]} \left(\text{sat}_{[-\rho, \beta - \rho]} \left(\frac{z_3}{z_2} \right) z_3 - \frac{\beta}{\rho} z_3 z_2 - \beta z_2^2 \right)$$

instead of $\psi(z)$, where $\text{sat}_{[\cdot]}$ denotes the saturation function

$$\text{sat}_{[a,b]}(x) = \max(a, \min(b, x)).$$

Finally, we choose the gains G_i of the observer such that $Sp(A + GC) = \{\lambda_1, \dots, \lambda_n\}$ with $\lambda_n < \lambda_{n-1} < \dots < \lambda_1 < 0$ and λ_1 enough negative. This amounts to take $G_i = \sigma_i(\{\lambda_1, \dots, \lambda_n\})$. One thus obtains the internal dynamics of the observer

$$(4.15) \quad \begin{cases} \dot{\hat{z}}_1 &= \hat{z}_2 + (\lambda_1 + \lambda_2 + \lambda_3)(\hat{z}_1 - y(t)) \\ \dot{\hat{z}}_2 &= \hat{z}_3 + (\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3)(\hat{z}_1 - y(t)) \\ \dot{\hat{z}}_3 &= \tilde{\psi}(\hat{z}) + \lambda_1 \lambda_2 \lambda_3 (\hat{z}_1 - y(t)) \end{cases}$$

and the estimators

$$(4.16) \quad \begin{cases} \hat{S} &= \frac{1}{\beta} \left(\text{sat}_{[-\rho, \beta-\rho]} \left(\frac{\hat{z}_3}{\hat{z}_2} \right) + \rho \right) \\ \hat{I} &= \frac{\hat{z}_2}{\rho} \\ \hat{R} &= N - \hat{S} - \hat{I} \end{cases} .$$

4.5. Discussion. The construction of an observer can avoid in certain situations to study the identifiability. For instance, in Sections 4.2 and 4.3, the observers do not require the knowledge of all the parameters of the model, and even of some functions involved in the model. This is why such observers are also called *unknown-inputs observers*. The theory of unknown-inputs observers has been mainly studied for linear systems [99, 39, 139]. Very few general results are available for nonlinear systems (this research field is today largely open).

The existence of observers without the possibility of fixing arbitrarily the speed of convergence, as in Examples 4.3 and 4.5, is connected to the property of *detectability* (see for instance [135, 9]), which is a weaker than observability: a system

$$(S) : \begin{cases} \dot{x} = f(x) \\ y = h(x) \end{cases}$$

is *detectable* if for any pair of solutions $x^a(\cdot)$, $x^b(\cdot)$ of (S), one has

$$\{h(x^a(t)) = h(x^b(t)), t \geq 0\} \Rightarrow \lim_{t \rightarrow +\infty} x^a(t) - x^b(t) = 0.$$

For linear dynamics $\dot{x} = Ax$, $y = Cx$ that are not observable, there exists a *Kalman decomposition* [109] i.e. an invertible matrix P such that

$$PAP^{-1} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}, \quad CP^{-1} = [C_1 \quad 0]$$

with $A_{11} \in \mathcal{M}_{l \times l}$, $C_1 \in \mathcal{M}_{1 \times l}$ where $l < n$ is the rank of the observability matrix \mathcal{O} recalled in (2.4), such that the subsystem $\dot{z} = A_{11}z$, $y = C_1z$ is observable. Then, the system is detectable when the matrix A_{22} is Hurwitz. This is exactly the case of Example 4.3 with $l = 1$.

5. Practical and numerical considerations.

5.1. Practical identifiability. Till now we have studied structural identifiability/observability. While structural identifiability is a property of the model structure, given a set of outputs, practical identifiability is related to the actual data. In particular, it is a measure of the amount of information contained in those 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 [155]. Structural identifiability means that parameters are identifiable with ideal (continuous, noise-free) data. While structural identifiability is a prerequisite for parameter identification, it does not guarantee that parameters are practically identifiable with a finite number of noisy data points.

Moreover, parameter estimation requires using numerical optimization algorithms. The distance, for the problem considered, to the nearest ill-posed problem, [60, 96], i.e., the conditioning of the problem, can challenge the convergence of algorithms. Another source of practical unidentifiability is the lack of information from the data, i.e., the signal from the data does not satisfy the persistence of excitation [121]. This is the case when the observation is near an equilibrium [19].

In this section we use sensitivity analysis and results from asymptotic statistical theory to study practical identifiability. We refer to previous surveys and papers on the topic [45, 17, 18, 55, 20]. Our purpose here is to give an intuitive account of these techniques.

5.1.1. Rationale for using sensitivity analysis. Practical identifiability is often assessed in terms of confidence intervals on parameters [203]. Confidence intervals can be derived from the Fisher Information Matrix (FIM) [27]. More specifically, the covariance matrix (Σ) of the estimated parameters may be approximated as the inverse of the FIM. The diagonal elements of $\Sigma \approx \text{FIM}^{-1}$ correspond to the variance of the parameter estimates. Their square-roots (the standard deviations) give confidence intervals on the parameters, thus providing information on practical identifiability.

In the least-squares framework, the Fisher Information Matrix can be expressed in terms of sensitivity matrices, that we define below.

5.1.2. Observed system. We consider that the initial condition x_0 is unknown. Unless otherwise specified, the term “parameter” now refers to both the parameter θ and the initial condition x_0 , i.e. $\Theta = (\theta, x_0)$. We make explicit the dependence of the state variables x and y on Θ to clarify the following derivations.

$$(5.1) \quad \begin{cases} \dot{x}(t, \Theta) = f(x(t, \Theta), \theta), & x(0, \Theta) = x_0, \\ \dot{\theta}(t) = 0, & \theta(0) = \theta, \\ y(t, \Theta) = h(x(t, \Theta), \theta). \end{cases}$$

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

5.1.3. Sensitivity analysis. We wish to quantify how the observed variable $y(t, \Theta)$ changes for a small parameter variation $\Delta\Theta$.

We denote the Jacobian of the observation $y(t, \Theta)$ with respect to the parameter Θ as

$$\chi(t, \Theta) = \frac{\partial y}{\partial \Theta}(t, \Theta).$$

This $m \times (n + p)$ matrix is called the sensitivity matrix.

By linearization (first-order Taylor approximation), one can write

$$\Delta y(t, \Theta) = \chi(t, \Theta) \Delta\Theta.$$

Side remarks. Reid [158] defined a parameter vector as “sensitivity identifiable” if the above equation can be solved uniquely for $\Delta\Theta$. This linear problem is well known: if χ has maximal rank then the solution is given by means of the Moore-Penrose pseudo-inverse $\chi^+ = (\chi^T \chi)^{-1} \chi^T$:

$$\Delta\Theta = \chi^+(t, \Theta) \Delta y(t, \Theta).$$

It is also well known [86] that the sensitivity of this solution is ruled by the condition number $\kappa_2(\chi) = \sigma_{\max}/\sigma_{\min}$ with σ_{\max} and σ_{\min} respectively the greatest and smallest singular value of χ (which are the corresponding eigenvalues of $\chi^T \chi$).

5.1.4. Ordinary Least Squares. Now we consider a set of M observations Y_i , $i = 1, \dots, M$, that have been obtained at times t_i . We assume that the observation is given by

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

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 constant variance $\text{var}(\mathcal{E}_i) = \sigma^2$;
- the errors are independent and identically distributed.

The Fisher Information Matrix, for the preceding defined observations, is defined as

$$(5.2) \quad \text{FIM}(\Theta, \sigma) = \frac{1}{\sigma^2} \sum_{i=1}^M \chi(t_i, \Theta)^T \chi(t_i, \Theta).$$

Solving the ordinary least square (OLS) equations gives an estimator $\hat{\Theta}_{\text{OLS}}$ of the parameter Θ :

$$(5.3) \quad \hat{\Theta}_{\text{OLS}} = \arg \min_{\Theta} \sum_{i=1}^M [Y_i - y(t_i, \Theta)]^2$$

Even though the error's distribution is not specified, asymptotic statistical theory can be used to approximate the mean and variance of the estimated Θ (a random variable) [170, 20]: the bias-adjusted approximation for σ^2 (with $n + p$ "parameters") is

$$(5.4) \quad \hat{\sigma}_{\text{OLS}}^2 = \frac{1}{M - (n + p)} \sum_{i=1}^M [Y_i - y(t_i, \hat{\Theta}_{\text{OLS}})]^2.$$

5.1.5. Confidence intervals. The above approximation of the error variance can be used to further approximate the parameter covariance matrix Σ :

$$(5.5) \quad \hat{\Sigma} := \left[\text{FIM} \left(\hat{\Theta}, \hat{\sigma}_{\text{OLS}}^2 \right) \right]^{-1}.$$

The standard error (SE) for $\hat{\Theta}_{\text{OLS}}$ can be approximated by taking the square roots of the diagonal elements of the covariance matrix Σ : for all $k = 1, \dots, n + p$,

$$(5.6) \quad \text{SE} \left(\hat{\Theta}_{\text{OLS}}(k) \right) = \sqrt{\hat{\Sigma}_{kk}}.$$

Finally, to compute the 95% confidence interval for the k -th component of the parameter vector $\hat{\Theta}_{\text{OLS}}$ with $n + p$ "parameters", one may use the Student's t -distribution with $M - (n + p)$ degrees of freedom: letting

$$\zeta(k) = t_{0.025}^{M-(n+p)} \times \text{SE} \left(\hat{\Theta}_{\text{OLS}}(k) \right),$$

the confidence interval is defined as

$$\hat{\Theta}_{\text{OLS}}(k) - \zeta(k) < \hat{\Theta}_{\text{OLS}}(k) < \hat{\Theta}_{\text{OLS}}(k) + \zeta(k).$$

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.

5.1.6. Computing the sensitivity matrix. The sensitivity matrix $\chi(t, \Theta)$, with $\Theta = (\theta, x_0)$, is obtained by integrating an ODE. The components of the ODE to be integrated depend on whether one differentiates with respect to θ or x_0 .

Differentiating with respect to θ . The first part of the ODE is given by

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

The Jacobian $\frac{\partial h}{\partial x}$ is a $m \times n$ matrix while $\frac{\partial h}{\partial \theta}$ is a $m \times p$ matrix.

We then have to compute the $n \times p$ matrix

$$z(t, \Theta) := \frac{\partial x}{\partial \theta}(t, \theta, x_0) = \frac{\partial x}{\partial \theta}(t, \Theta).$$

Let $A(t, \Theta)$ and $B(t, \Theta)$ be the following time-dependent $n \times n$ and $n \times p$ matrices, respectively:

$$A(t, \Theta) := \frac{\partial f}{\partial x}(x(t, \Theta), \theta),$$

and

$$B(t, \Theta) := \frac{\partial f}{\partial \theta}(x(t, \Theta), \theta).$$

It is well known [93] that $z(t)$ is the solution of the linear matrix equation:

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

with the initial condition $z(0, \Theta) = 0_{n \times p}$ (a zero matrix of size $n \times p$).

Differentiating with respect to x_0 . The second part of the ODE is given by

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

Let

$$w(t, \Theta) := \frac{\partial x}{\partial x_0}(t, \theta, x_0) = \frac{\partial x}{\partial x_0}(t, \Theta).$$

Based on the same reference [93], $w(t, \Theta)$ is solution of the linear matrix ODE

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

with the initial condition $w(0, \Theta) = \text{Id}_{n \times n}$ (the identity matrix of size n).

Full system. To summarize, one has to solve the following system in dimension $n^2 + np + n$

$$(5.7) \quad \begin{cases} \dot{x}(t, \Theta) = f(x(t, \Theta), \theta), & x(0, \Theta) = x_0, \\ \dot{z}(t, \Theta) = A(t, \Theta) z(t, \Theta) + B(t, \Theta), & z(0, \Theta) = 0_{n \times p}, \\ \dot{w}(t, \Theta) = A(t, \Theta) w(t, \Theta), & w(0, \Theta) = \text{Id}_{n \times n}, \end{cases}$$

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

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

5.1.7. Some case studies. In this section, we consider two classical examples as case studies. These examples have been used in many books of mathematical epidemiology, e.g. [137].

Case 1. Influenza in a boarding school. Our first example is an outbreak of influenza in a United Kingdom boarding school which occurred in 1978 [38]. In [137] the parameters β, γ are identified by an unspecified “best-fit” algorithm. A more complete analysis is done in [30], where the analysis is done using sensitivity analysis and asymptotic statistical theory. In [127] the same example is considered. Different sources exist for the data [31, 30, 59] with small differences.

Using the figure in [38] and the *Plot Digitizer* software, we got an approximation of the data. It was reported that $N = 763$, and the conditions at the start of this outbreak were $S_0 = 762$ and $I_0 = 1$. We used the following data, in which time t is in day and $i(t)$ denotes the number of infectious people at time t .

t	0	1	2	3	4	5	6	7	8	9	10	11	12	13
$i(t)$	1	6	26	73	222	293	258	237	191	124	68	26	10	3

Specifically, we considered model (3.4) with $k = 1$ (all infectious are assumed to be observed). We obtained the following OLS estimation, as given by the *Scilab* software:

$$\beta \approx 1.9605032, \quad \gamma \approx 0.4751562$$

(see the numerical code in Appendix C). The fit is shown in Figure 1.

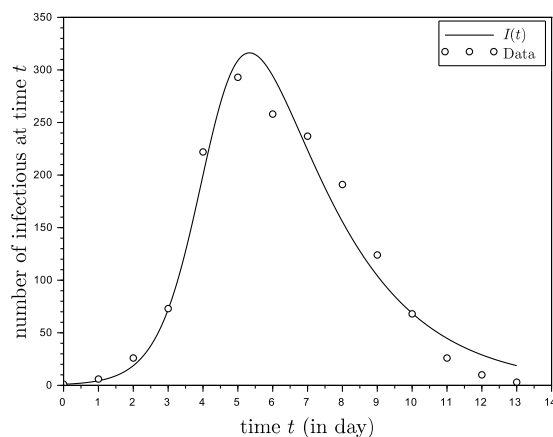


FIGURE 1. *Boarding School example.*

We computed the 95% confidence intervals using the formulas given in the preceding section: (5.2), (5.3), (5.4), (5.5), (5.6), up to a few changes due to the fact that the initial conditions $x_0 = (S_0, I_0)$ are assumed to be known in this example (see Appendix C). We find:

$$\beta \approx 1.9605032 \pm 0.0731602,$$

$$\gamma \approx 0.4751562 \pm 0.0408077.$$

One can obtain approximately the same results using the likelihood profile method to compute confidence intervals [27], assuming normally distributed errors. However, it is well known that the profile method quickly becomes impractical for model with more than two parameters [27], which is the rule rather than the exception, as will be the case in the following example. This is why we stick to the FIM method. Lastly, we note that the condition number of the FIM is approximately equal to 3.80 in this example.

Case 2. Plague in Bombay. Our second example is the Bombay Plague of 1905–1906 [112]. We collected the data from [53, Table IX], over the same period as [112] (Dec. 17 to Jul. 21). The form of the data is presented in the following table, in which time t is in week, and $\dot{r}(t)$ denotes the number of death per week at time t .

t	0	1	2	3	4	5	6	...	24	25	26	27	28	29	30
$\dot{r}(t)$	8	10	12	16	24	48	51	...	106	64	46	35	27	28	24

We consider that the number of death per week is the same as $\dot{R}(t) = \gamma I(t)$, meaning that all infections lead to death, which is a reasonable assumption in this context [14]. Therefore, we consider model (3.4) with $k = \gamma$. In this example, not only the parameters β and γ , but also the size of the population, N , as well as the initial conditions, S_0 and I_0 , are unknown [14]. According to Theorem 3.1 (and Remark 3.2), the model is neither observable nor identifiable. However, the model is partly identifiable in the sense that S_0 , I_0 , γ and $\tilde{\beta} = \beta/N$ are structurally identifiable. Starting from an arbitrary initial guess, we obtained the following OLS estimation, as given by the `Scilab` software:

$$\tilde{\beta} \approx 0.0000855, \quad \gamma \approx 3.7161743, \quad S_0 \approx 48113.13, \quad I_0 \approx 1.4213612,$$

(see the numerical code in Appendix D). The fit is shown in Figure 2.

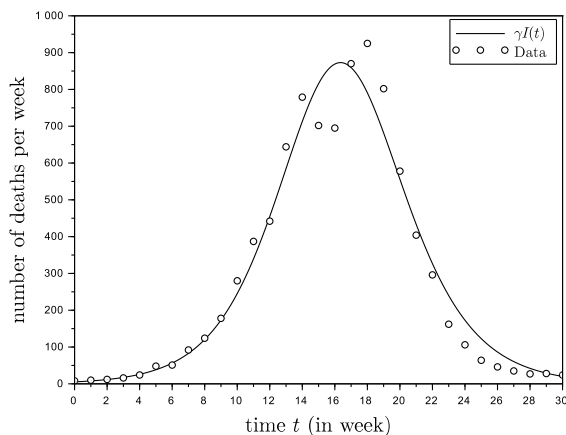


FIGURE 2. *Plague in Bombay example.*

Proceeding as in the previous example, we obtained the following 95%-confidence

intervals:

$$\tilde{\beta} \approx 0.0000855 \pm 0.0015784,$$

$$\gamma \approx 3.7161743 \pm 25.255243,$$

$$S_0 \approx 48113.13 \pm 593794.26,$$

$$I_0 \approx 1.4213612 \pm 34.439929.$$

The confidence intervals are huge, which means that we can have absolutely no confidence in the estimated values of the parameters, even though the fit looks good and these parameters are structurally identifiable in principle. In practice, one can show that many other and very different combinations of the parameters can yield approximately the same fit. Note that if we did as if the initial conditions were known, the confidence intervals on $\tilde{\beta}$ and γ would be reasonable, as in the previous example. In this example, the condition number of the FIM is approximately 9.14×10^{24} , meaning that the problem is “sloppy” [42]; note however that while it is usual that a model is both sloppy and practically non-identifiable, this is not always the case [52]. Altogether, we can conclude that there is a severe practical identifiability issue in this classical example.

5.1.8. Discussion. The SIR model of Kermack-McKendrick has been studied in a series of papers [46, 45, 47, 35] where the problem of observability/identifiability is approached from the statistical point view: addressing parameter identifiability by exploiting properties of both the sensitivity matrix and uncertainty quantification in the form of standard errors. In this series of papers, structural observability and identifiability were not explicitly addressed. For example in [47] the authors identify $(S_0, I_0, \beta/N, \gamma)$ based on incidence observations - akin to equation (3.5) with $k = 1$ - which we have proved to be structurally identifiable (see Theorem 3.3). Similarly in [35] the authors seek to identify $(S_0, I_0, \beta, \gamma)$ which, with N known, are structurally identifiable in principle. However the authors encounter practical identifiability issues. This is a typical example of strictly practical unidentifiability (as in our second example, the Plague in Bombay).

Although a structural observability and identifiability analysis should be done as a prerequisite to a practical identifiability analysis, it does not suffice. Moreover, when doing practical identifiability analyses, the error structure of the data should be considered. For instance, sensitivity analyses can be extended to non-constant error variance through Generalized Least Squares (GLS), which makes it possible to test different ways of weighting errors [46]. An additional issue may occur when the output signal is not sufficiently informative (i.e., not *persistently exciting* [121]). For example when the data correspond to states near unobservability, e.g., near an equilibrium. In those cases, one has to wait to have data sufficiently far from equilibrium.

To conclude, the problem of observability and identifiability, either structural or practical, is far from being simple, even in relatively simple SIR models with seemingly good quality data [46]. Of course, the more complex the model, the more parameters there are to identify, the more serious the problem of identifiability.

5.2. Observers in practice. In this section, we show how the various observers presented in Section 4 behave in practice, and the role of the tuning parameters. Up to now, we have assumed the measurements to be perfect i.e. not tainted with any noise. Since integration has good “averaging” properties, an observer is expected to filter noise or inaccuracies in the measurements. However, we will see that the

filtering capacity of an observer is related to his convergence speed, which often leads to a “precision-speed” dilemma in the choice of the observer or his settings.

Let us underline that when identifiability/observability cannot be proved theoretically or is too difficult to be proven analytically, one can still look for an observer and study its asymptotic convergence, theoretically or numerically.

5.2.1. Observers with linear assignable error dynamics. We illustrate the observer (4.6) of the age-structured model (4.5) on simulations, for the following values of the parameters.

a_1	a_2	m_1	m_2	m_3	k	\bar{r}_{min}	\bar{r}_{max}
0.1	0.1	0.05	0.07	0.07	1	0.9	1.1

The following code is used to compute the gain vector \mathbf{G} for a set of desired eigenvalues.

```
a1=0.1;a2=0.1;m1=0.05;m2=0.07;m3=0.07;
Sp=0.3*[-0.3,-0.33,-0.36];
A=[-a1-m1,0,0;
    a1,-a2-m2,0;
    0,a2,-m3];
C=[0,0,1];
B=[0;0;1];
Obs=[C;C*A;C*A*A];
L=inv(Obs)*B;
P=[L,A*L,A*A*L];
Abar=inv(P)*A*P;
sigma=coeff(poly(Sp,'x'));
G=P*(-sigma(1:3)'+Abar(:,3));
```

Figure 3 shows convergence for a moderately negative spectrum, while Figure 4 shows the acceleration of convergence obtained for a spectrum located further to the left in the complex plane. For the same choice of gains, Figures 5 and 6 show the effect of noise on the $y(\cdot)$ measurements. It can be seen that a faster convergence is more sensitive to noise and loses accuracy. In practice, one often has to make a compromise for the choice of the observer’s setting.

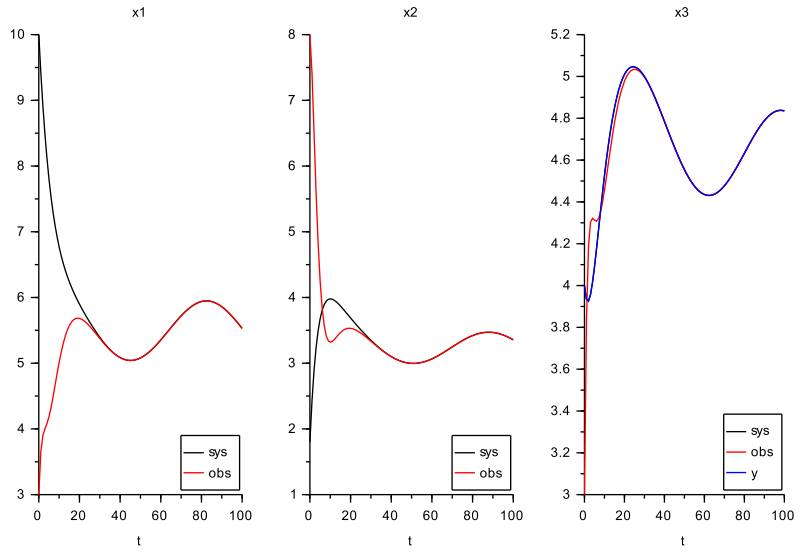


FIGURE 3. $Sp(A + GC) = \{-0.3, -0.33, -0.36\}$ without measurement noise

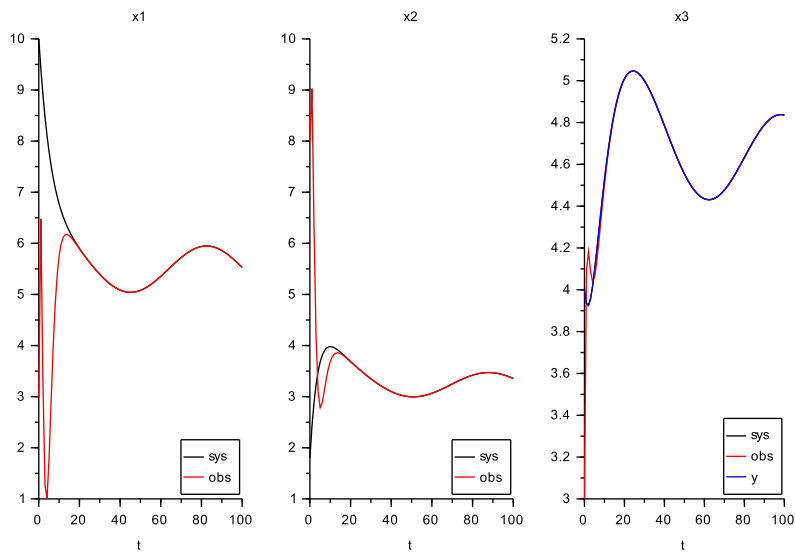


FIGURE 4. $Sp(A + GC) = \{-0.6, -0.66, -0.72\}$ without measurement noise

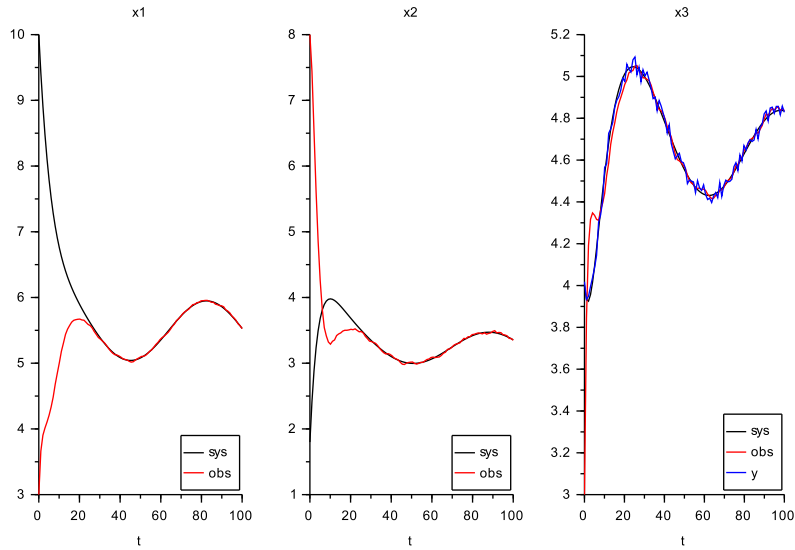


FIGURE 5. $Sp(A + GC) = \{-0.3, -0.33, -0.36\}$ with measurement noise

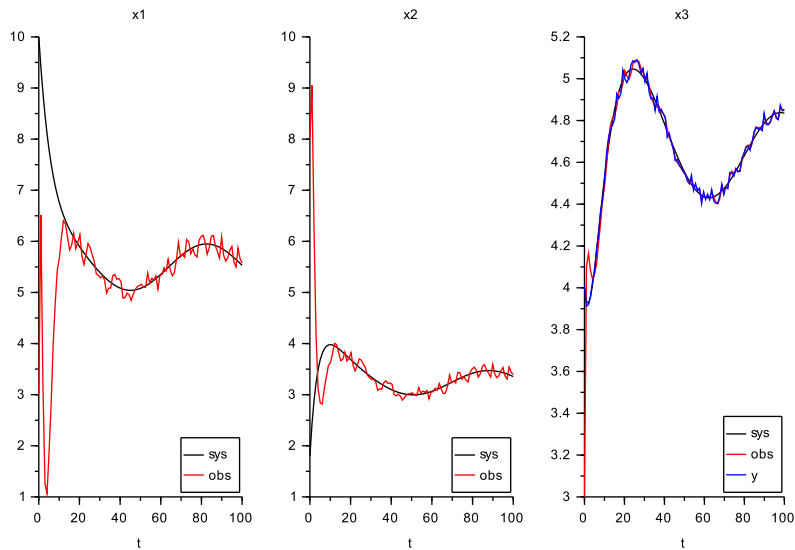


FIGURE 6. $Sp(A + GC) = \{-0.6, -0.66, -0.72\}$ with measurement noise

5.2.2. Observers with partially assignable error dynamics. The observer given in (4.8) for the intra-host malaria model (2.11) is illustrated here on real data, as one can see for instance on Figure 7.

As already mentioned in Remark 4.4, the speed of convergence of this observer

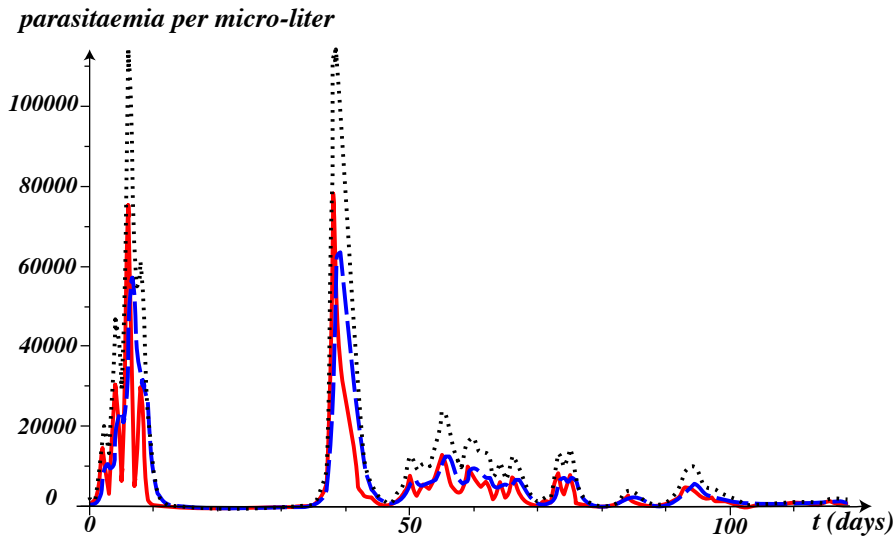


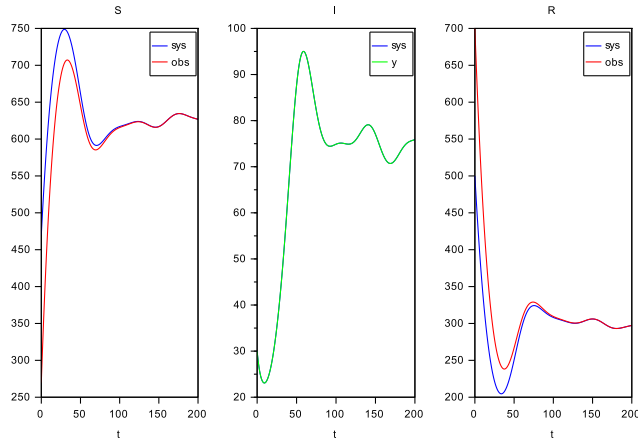
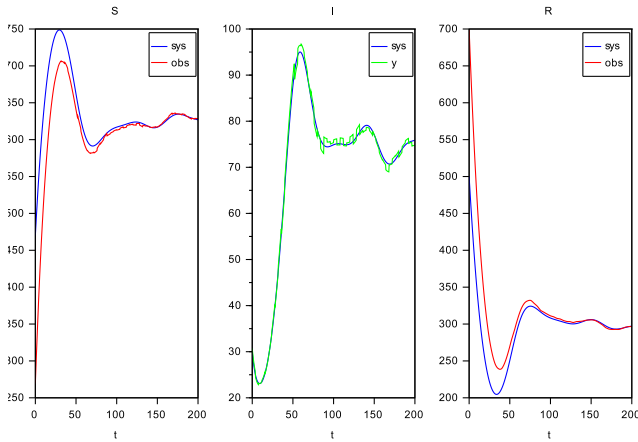
FIGURE 7. Example of patient S1204: Measures (data) of peripheral parasitaemia are plotted with red solid line, The estimations delivered by the observer (4.8) are plotted with blue dashed line for the estimated sequestered parasitaemia, and with black dotted line for the estimated total parasitaemia. The gain used is $L = (0, 5, 5, 0, 0, 0, 0)^T$. $Sp(\bar{A} - LC) = \{-10, -\mu_S, -\mu_M, -\mu_5 - \gamma_5, -\mu_4 - \gamma_4, -\mu_3 - \gamma_3, -\mu_2 - \gamma_2 - \gamma_1\}$.

cannot be tuned as fast as desired. However, this is quite satisfactory in practice. Let us also underline that the observer does not require the reconstruction of the parameter β , although this parameter is identifiable (see Section 2.4). This is a strength of this observer, because the parameter β could switch or fluctuate with time.

5.2.3. Observers with asymptotic convergence. We illustrate on simulations the behavior of the asymptotic observer (4.10) of the SIR model with fluctuating rates (4.9), for the following values of the parameters.

β	ρ	ν	μ	N
0.4 ± 0.08	0.2 ± 0.04	0.05	0.05	1000

Here β and ρ are functions of time chosen randomly in between the bounds given in the table. Figures 8 and 9 show that the observer has a convergence relatively insensitive to measurement noise, but the speed of convergence is slow because the exponential decay of the error is equal to μ , which is not adjustable.

FIGURE 8. *without measurement noise*FIGURE 9. *with measurement noise*

Unlike the observers in previous sections, let us underline that the present observer is not based on a $\hat{y} - y$ innovation. Therefore, one is not informed of the quality of the estimate over time, which is a *price to pay* to have an observer insensitive to unknown variations of the epidemic parameters β, ρ .

5.2.4. High gain observer. The non-linear observer (4.15)-(4.15) of the classical SIR model (4.14) is illustrated on simulations for the following values

β	ρ	N
0.4	0.1	10000

where the y cumulative measures were made discretely every day (rounded to the nearest integer). In order to obtain a time-continuous $y(\cdot)$ signal, we performed an interpolation by cubic splines. Figure 10 shows the convergence of the observer for the eigenvalues $\{-2, -2.2, -2.4\}$.

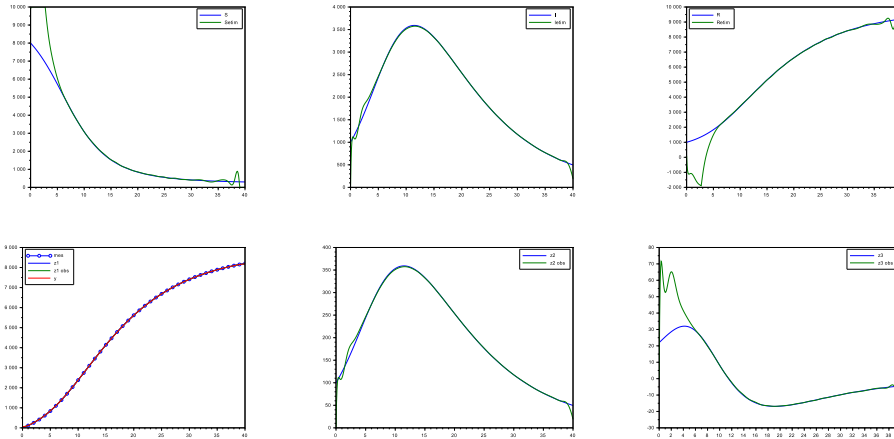


FIGURE 10. Observer simulations (top: variables S , I , R and their estimates; bottom: coordinates z_i with measurements points on the left)

We also simulated the observer when the measurements are corrupted by random counting errors up to ± 5 individuals per day (see Figure 11).

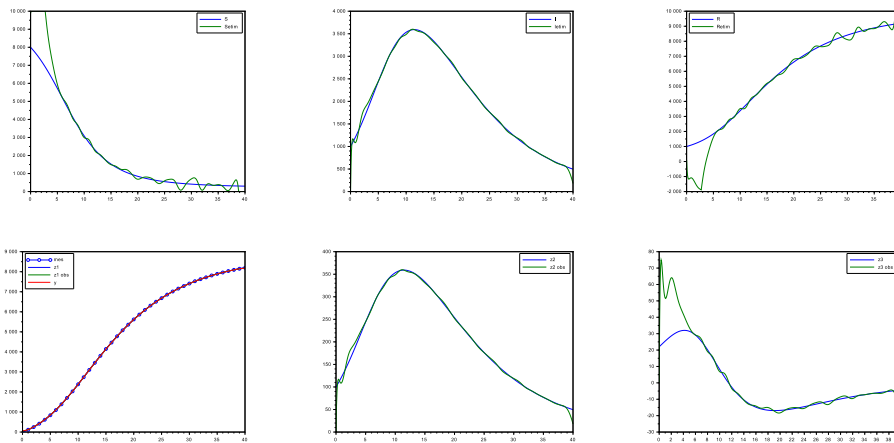


FIGURE 11. Observer simulations with noisy measurements

As for the adjustable observer in Section 5.2.1, these simulations show the dilemma **accuracy versus speed** of the estimation in presence of measurement noise.

Appendix A. Proof of Lemma 4.1.

The proof is adapted from [8].

Let us first consider pairs (\bar{A}, \bar{C}) of the canonical form known as Brunovsky's form

$$\bar{A} = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & \cdots & 0 & -a_{n-1} \\ \cdot & \ddots & \ddots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \ddots & \ddots & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & 0 & -a_2 \\ 0 & \cdots & \cdots & 0 & 1 & -a_1 \end{bmatrix}, \quad \bar{C} = [0 \quad \cdots 0 \quad 1]$$

where the a_i are any numbers. Their observability matrices are lower triangular:

$$\bar{O} = \begin{bmatrix} & & & & 1 \\ & 0 & & \ddots & \\ & & \ddots & & \\ & & & \star & \\ 1 & & & & \end{bmatrix}$$

therefore invertible. It is easy to see that the characteristic polynomial of the matrix \bar{A} is given by

$$\pi_{\bar{A}}(\xi) = \xi^n + a_1 \xi^{n-1} + \cdots + a_{n-1} \xi + a_n$$

Indeed, if X is a left eigenvector of \bar{A} for an eigenvalue λ (possibly complex), $X\bar{A} = \lambda X$ gives

$$\begin{aligned} X_2 &= \lambda X_1, \\ X_3 &= \lambda X_2 = \lambda^2 X_1, \\ &\vdots \\ X_n &= \lambda X_{n-1} = \lambda^{n-1} X_1, \\ -a_n X_1 - a_{n-1} X_2 - \cdots - a_1 X_n &= \lambda X_n. \end{aligned}$$

Thus the line vector X is of the form

$$X = [1 \quad \lambda \quad \lambda^2 \quad \cdots \quad \lambda^{n-1}] X_1 \quad \text{with } X_1 \neq 0$$

and λ verifies

$$(\lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + \cdots + a_{n-1} \lambda + a_n) X_1 = 0.$$

Since X_1 is non-zero, we deduce that the eigenvalues are roots of the polynomial

$$\lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + \cdots + a_{n-1} \lambda + a_n = 0$$

which is of degree n and whose coefficient of λ^n is equal to 1.

The characteristic polynomial of the matrix $\bar{A} + \bar{G}\bar{C}$, where \bar{G} is a vector of \mathbb{R}^n with elements denoted \bar{g}_i , is written as follows

$$\pi_{\bar{A} + \bar{G}\bar{C}}(\xi) = \xi^n + (a_1 - \bar{g}_n) \xi^{n-1} + \cdots + (a_{n-1} - \bar{g}_2) \xi + (a_n - \bar{g}_1)$$

Thus, one can arbitrarily choose the n coefficients of this polynomial by choosing the n elements of \bar{G} , and thus freely assign the spectrum of the matrix $\bar{A} + \bar{G}\bar{C}$. For any set $\Lambda = \{\lambda_1, \cdots, \lambda_n\}$ of n real or complex numbers two by two conjugates, one has just to identify the coefficients of the polynomial $\pi_{\bar{A} + \bar{G}\bar{C}}$ with those of

$$\prod_{i=1}^n (\xi - \lambda_i) = \xi^n + \sum_{k=1}^n (-1)^k \sigma_k(\Lambda) \xi^{n-k}$$

Thus, we obtain

$$\bar{g}_i = a_{n+1-i} + (-1)^{n-i} \sigma_{n+1-i}(\Lambda), \quad i = 1 \cdots n$$

Let us now show that for any pair (A, C) such that O is full rank, there is an invertible P matrix such that $P^{-1}AP = \bar{A}$ and $CP = \bar{C}$, where the pair (\bar{A}, \bar{C}) is in the Brunovsky's form. Consider the vector

$$L = O^{-1} \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{bmatrix} \Rightarrow \begin{cases} CA^k L = 0, & k = 0 \cdots n-2 \\ CA^{n-1} L = 1 \end{cases}$$

and the matrix consisting of the concatenation of the columns

$$P = [L \ AL \ \cdots \ A^{n-1}L]$$

We have

$$OL = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{bmatrix}, \quad OAL = \begin{bmatrix} CAL \\ CA^2L \\ \vdots \\ CA^{n-1}L \\ CA^nL \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ 1 \\ \star \end{bmatrix}, \quad OA^2L = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \star \\ \star \end{bmatrix}, \cdots$$

up to

$$OA^{n-1}L = \begin{bmatrix} 1 \\ \star \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix}$$

Thus the OP matrix is of the form

$$OP = \begin{bmatrix} & & & 1 \\ & 0 & \ddots & \\ & \ddots & \star & \\ 1 & & & \end{bmatrix}$$

which shows that P is indeed an invertible matrix. Finally, the columns of the AP matrix are

$$AP = [AL \ A^2L \ \cdots \ A^{n-1}L \ A^nL]$$

Its first $n-1$ columns are written as follows

$$[AL \ A^2L \ \cdots \ A^{n-1}L] = P \begin{bmatrix} 0 & \cdots & & \\ 1 & 0 & \cdots & \\ 0 & \ddots & \ddots & \\ \vdots & & & 1 & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

By Cayley-Hamilton's Theorem, we have $\pi_A(A) = 0$, which allows us to write the last column of AP as

$$A^n L = -a_n L - a_{n-1} AL - \dots - a_1 A^{n-1} L = P \begin{bmatrix} -a_n \\ -a_{n-1} \\ \vdots \\ -a_1 \end{bmatrix}$$

which shows that we have $PA = \bar{A}P$. Finally, one gets

$$CP = [CL \quad CAL \quad \dots \quad CA^{n-1}L] = [0 \quad \dots \quad 0 \quad 1] = \bar{C}$$

Then, for a vector \bar{G} such that $Sp(\bar{A} + \bar{G}\bar{C}) = \Lambda$, we have $Sp(P^{-1}(\bar{A} + \bar{G}\bar{C})P) = \Lambda$, where $P^{-1}(\bar{A} + \bar{G}\bar{C})P = A + P\bar{G}C$. We conclude that for the gain vector $G = P\bar{G}$, we have $Sp(A + GC) = \Lambda$.

Appendix B. Proof of Lemma 4.6.

The proof is adapted from [76].

Let X be a left eigenvector of $A + GC$ for the eigenvalue λ_i . By writing $X(A + GC) = \lambda_i X$, we obtain the $n - 1$ inequalities.

$$\begin{aligned} X_1 &= \lambda_i X_2 \\ X_2 &= \lambda_i X_3 \\ &\vdots \\ X_{n-1} &= \lambda_i X_n \end{aligned}$$

Thus X_n is necessarily non-zero and can be taken equal to 1, which gives

$$X = [\lambda_i^{n-1} \quad \lambda_i^{n-2} \quad \dots \quad \lambda_i \quad 1]$$

We then obtain the n rows of the matrix $V_{\lambda_1, \dots, \lambda_n}$, which defines a matrix of change of basis that diagonalizes the matrix $A + GC$.

Now let's show how to determine the inverse of $V_{\lambda_1, \dots, \lambda_n}$. Let w_{ij} be the coefficients of $V_{\lambda_1, \dots, \lambda_n}^{-1}$. The equality $(V_{\lambda_1, \dots, \lambda_n}^{-1})(V_{\lambda_1, \dots, \lambda_n}) = Id$ gives

$$(B.1) \quad \sum_{k=1}^n w_{kj} \lambda_i^{nk} = \delta_{ij} := \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

For each j in $\{1, \dots, n\}$, let's consider the polynomial

$$(B.2) \quad P_j(X) = \sum_{k=1}^n w_{kj} X^{n-k}$$

The conditions (B.1) amount to write $P_j(\lambda_i) = \delta_{ij}$, i.e. the polynomial P_j has $n - 1$ roots λ_i for $i \neq j$ and $P_j(\lambda_j)$ is equal to 1. So it has the following expression

$$P_j(X) = \prod_{k \neq j} \frac{X - \lambda_k}{\lambda_j - \lambda_k}$$

By identifying its coefficients with those of the expression (B.2), we obtain

$$(B.3) \quad w_{ij} = (-1)^{i-1} \frac{\sigma_{i-1}(\Lambda \setminus \{\lambda_j\})}{\prod_{k \neq j} \lambda_j - \lambda_k}$$

where the σ_k are the symmetric functions defined in (4.4).

Let

$$\varphi(\lambda_1, \dots, \lambda_n) = \lambda_1 + C \|V_{\lambda_1, \dots, \lambda_n}^{-1}\|_{\infty} + \theta$$

The expression (B.3) shows that the norm $\|V_{\lambda_1, \dots, \lambda_n}^{-1}\|_{\infty}$ becomes arbitrarily large when $\lambda_i - \lambda_j$ approaches 0 (for $i \neq j$), which ensures the existence of numbers $\lambda_n < \lambda_{n-1} < \dots < \lambda_1 < 0$ such as $\varphi(\lambda_1, \dots, \lambda_n) > 0$. For $\lambda_i = -\alpha^i$ ($i = 1, \dots, n$), we obtain, for any j

$$\lim_{\alpha \rightarrow +\infty} w_{ij} = \begin{cases} 0 & i < n \\ 1 & i = n \end{cases}$$

and $\|V_{-\alpha, -\alpha^2, \dots, -\alpha^n}^{-1}\|_{\infty}$ thus tends towards 1 when α tends towards $+\infty$, which shows the existence of numbers $\lambda_n < \lambda_{n-1} < \dots < \lambda_1 < 0$ such that $\varphi(\lambda_1, \dots, \lambda_n) < 0$. Finally, by continuity of φ , we deduce the existence of $\lambda_n < \lambda_{n-1} < \dots < \lambda_1 < 0$ such that $\varphi(\lambda_1, \dots, \lambda_n) = 0$.

Appendix C. Implementation of the ‘‘Boarding School’’ example.

C.1. Derivation of the Fisher Information Matrix. In this example, we consider $x \in \mathbb{R}^n$ with $n = 2$, $y \in \mathbb{R}^m$ with $m = 1$, i.e.,

$$x(t) = \begin{pmatrix} S(t) \\ I(t) \end{pmatrix}, \quad y(t) = I(t),$$

and $\theta \in \mathbb{R}^p$ with $p = 2$, i.e.,

$$\theta = \begin{pmatrix} \beta \\ \gamma \end{pmatrix}.$$

We consider the following model, equivalent to model (3.4) with $k = 1$:

$$\begin{aligned} \dot{x} = f(x, \theta) &= \begin{pmatrix} \dot{S} \\ \dot{I} \end{pmatrix} = \begin{pmatrix} f_S(S, I, \theta) \\ f_I(S, I, \theta) \end{pmatrix} = \begin{pmatrix} -\beta SI/N \\ \beta SI/N - \gamma I \end{pmatrix}, \quad x(0) = x_0 = \begin{pmatrix} S_0 \\ I_0 \end{pmatrix}, \\ y = h(x, \theta) &= I. \end{aligned}$$

We disregard $\Theta = (\theta, x_0)$ since the initial conditions are assumed to be known in this example. The Jacobian of the observation with respect to the parameter θ is:

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

since in this example,

$$\frac{\partial h}{\partial \theta}(t) = 0.$$

This Jacobian has dimension $m \times p = 1 \times 2$. We have

$$\frac{\partial h}{\partial x} = \begin{pmatrix} \frac{\partial h}{\partial S} & \frac{\partial h}{\partial I} \end{pmatrix} = \begin{pmatrix} 0 & 1 \end{pmatrix},$$

and

$$z = \frac{\partial x}{\partial \theta} = \begin{pmatrix} \frac{\partial S}{\partial \beta} & \frac{\partial S}{\partial \gamma} \\ \frac{\partial I}{\partial \beta} & \frac{\partial I}{\partial \gamma} \end{pmatrix}.$$

This yields

$$\chi = \frac{\partial h}{\partial x} \frac{\partial x}{\partial \theta} = \begin{pmatrix} \frac{\partial I}{\partial \beta} & \frac{\partial I}{\partial \gamma} \end{pmatrix}.$$

Let $\{t_i\}$, $i = 0, 1, 2, \dots, M$, be the sampling times. Fisher's Information Matrix is defined as:

$$\text{FIM}(\theta, \sigma) = \frac{1}{\sigma^2} \sum_{i=1}^M \chi(t_i, \theta)^T \chi(t_i, \theta),$$

where σ^2 is defined as the sum of the squared error (SSE) divided with $M - p$ instead of $M - (n + p)$ as in equation (5.4), since the initial conditions are assumed to be known in this example.

Computing Fisher's Information Matrix. Let

$$A(t) = \frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f_S}{\partial S} & \frac{\partial f_S}{\partial I} \\ \frac{\partial f_I}{\partial S} & \frac{\partial f_I}{\partial I} \end{pmatrix} = \begin{pmatrix} -\beta I/N & -\beta S/N \\ \beta I/N & \beta S/N - \gamma \end{pmatrix},$$

and

$$B(t) = \frac{\partial f}{\partial \theta} = \begin{pmatrix} \frac{\partial f_S}{\partial \beta} & \frac{\partial f_S}{\partial \gamma} \\ \frac{\partial f_I}{\partial \beta} & \frac{\partial f_I}{\partial \gamma} \end{pmatrix} = \begin{pmatrix} -SI/N & 0 \\ SI/N & -I \end{pmatrix}.$$

The matrix z can be computed by numerically solving the following system of ODE's:

$$\begin{aligned} \dot{x} &= f(x, \theta), & x(0) &= x_0, \\ \dot{z} &= Az + B, & z(0) &= 0_{n \times p}, \end{aligned}$$

which is a subsystem of (5.7) since the initial conditions are assumed to be known in this example (i.e., we disregard w). In the following code, the entries of x and z are indexed in this way:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad z = \begin{pmatrix} z_3 & z_5 \\ z_4 & z_6 \end{pmatrix},$$

leading to

$$\chi = \begin{pmatrix} z_4 & z_6 \end{pmatrix}.$$

C.2. Numerical implementation. The code has been written with the `Scilab` language and executed under `SCILAB 6.0.02`. It consists in a function for identifying β and γ and using the `lsqrsolve` function which implements the Levenberg-Marquard algorithm to perform ordinary least squares. We could have chosen the `fminsearch` function which is an implementation of the Nelder-Mead algorithm, but this gives exactly the same results. For solving ODE's, `Scilab` uses the `lsoda` solver of `ODEPACK`. It automatically selects between non-stiff predictor-corrector Adams method and stiff Backward Differentiation Formula (BDF) method. It uses non-stiff method initially and dynamically monitors data in order to decide which method to use.

We define the following functions in the `Scilab` environment:

²<https://www.scilab.org/>

```

function [kguess_n,SSE]=identifKMK(OBS,T,kguess,N)
// kguess_n = [BETA;GAMMA]
t0=T(1);m=length(T);
[x,SSE]=lsqrsolve(kguess,errorKmcK,m,[1.d-8,1.d-8,1.d-5,1d9,0,100]);
nbparam=length(kguess)
kguess_n=x;
x0=[N-OBS(1);OBS(1)];
sol=ode(x0,t0,T,list(KmcK,x(1),x(2)));
kguess_n=kguess_n(:);
SSE=sum(SSE.^2)
xset("window",1)
sol1=ode(x0,t0,T(1):0.01:T($),list(KmcK,x(1),x(2)));
clf
plot((T(1):0.01:T($))',sol1(2,:))
plot(T',OBS,'ro')
endfunction

function y=errorKmcK(k,m)
x0=[N-OBS(1);OBS(1)];x0=x0(:);
BETA=k(1);GAMMA=k(2);
sol=ode(x0,t0,T,list(KmcK,BETA,GAMMA));
predic=sol(2,:);
predic=predic(:);
OBS=OBS(:);
y=OBS-predic;
endfunction

function xdot=KmcK(t,x,BETA,GAMMA,N)
xdot=[-BETA/N*x(2),0;BETA/N*x(2),-GAMMA]*x
endfunction

```

Then, the Scilab session goes like this

```

--> load('databoarding')
ans =
T
--> OBS=dataBSFlu

OBS =

column 1 to 8

1.   6.   26.   73.   222.   293.   258.   237.

column 9 to 14

191.   124.   68.   26.   10.   3.

--> M=length(OBS);

```

```

--> T=0:M-1;

--> N=763;

--> beta0=2;gamma0=0.5;param=[beta0,gamma0];

--> p=length(param);

--> [param,SSE]=identifKMK(OBS,T,param,N)

SSE =

4892.6472

param =

1.9605032
0.4751562

--> sigma2=SSE/(M-p)

sigma2 =

407.72060

--> BETA=param(1);GAMMA=param(2);

```

We then compute confidence intervals using the formulas (5.2), (5.3), (5.4), (5.5), (5.6).

```

function FIM=fimKmcK(x0,T,BETA,GAMMA,sigma2)
// Compute the sensitivity matrix
x0=x0(:);t0=T(1);
X0=[x0;0;0;0;0;0]
sol=ode(X0,t0,T,list(JKmcK,BETA,GAMMA));
M=sol([4,6],:);
FIM=M*M'./sigma2;
endfunction

function xdot=JKmcK(t,x,BETA,GAMMA)
xdot(1)=-BETA*x(1)*x(2)/N
xdot(2)=BETA*x(1)*x(2)/N-GAMMA*x(2)
xdot(3)=-BETA*x(2)*x(3)/N-BETA*x(1)*x(4)/N-x(1)*x(2)/N
xdot(4)=BETA*x(2)*x(3)/N+(BETA*x(1)/N-GAMMA)*x(4)+x(1)*x(2)/N
xdot(5)=-BETA*x(2)*x(5)/N-BETA*x(1)*x(6)/N
xdot(6)=BETA*x(2)*x(5)/N+(BETA*x(1)/N-GAMMA)*x(6)-x(2)
endfunction

```

Then the Scilab session is

```

--> x0=[N-OBS(1); OBS(1)];

```

```

--> FIM=fimKmcK(x0,T,BETA,GAMMA)
FIM =

    974.5073   -523.73985
   -523.73985   3132.2047

--> cond(FIM)
ans =

3.8082403

--> CovMAT=inv(FIM)
CovMAT =

    0.0011275   0.0001885
    0.0001885   0.0003508

--> t=cdfT("T",M-p,0.975,0.025)
t =

2.1788128

--> confBETA=t*sqrt(CovMAT(1,1))
confBETA =

0.0731602

--> confGAMMA=t*sqrt(CovMAT(2,2))
confGAMMA =

0.0408077

```

Appendix D. Implementation of the “Plague in Bombay” example.

D.1. Derivation of the Fisher Information Matrix. In this example, we consider $x \in \mathbb{R}^n$ with $n = 2$, $y \in \mathbb{R}^m$ with $m = 1$, i.e.,

$$x(t) = \begin{pmatrix} S(t) \\ I(t) \end{pmatrix}, \quad y(t) = I(t),$$

and $\theta \in \mathbb{R}^p$ with $p = 2$, i.e.,

$$\tilde{\beta} = \frac{\beta}{N}, \quad \theta = \begin{pmatrix} \tilde{\beta} \\ \gamma \end{pmatrix}.$$

We consider the following model, equivalent to model (3.4) with $k = \gamma$:

$$\begin{aligned} \dot{x} = f(x, \theta) &= \begin{pmatrix} \dot{S} \\ \dot{I} \end{pmatrix} = \begin{pmatrix} f_S(S, I, \theta) \\ f_I(S, I, \theta) \end{pmatrix} = \begin{pmatrix} -\tilde{\beta}SI \\ \tilde{\beta}SI - \gamma I \end{pmatrix}, \quad x(0) = x_0 = \begin{pmatrix} S_0 \\ I_0 \end{pmatrix}, \\ y = h(x, \theta) &= \gamma I. \end{aligned}$$

We consider $\Theta = (\theta, x_0)$ since the initial conditions are assumed to be unknown in this example. The Jacobian of the observation with respect to the parameter Θ is:

$$\chi(t, \Theta) = \frac{\partial y}{\partial \Theta}(t) = \frac{\partial h}{\partial x} \frac{\partial x}{\partial \Theta}(t) + \frac{\partial h}{\partial \Theta}(t),$$

This Jacobian has dimension $m \times (p + n) = 1 \times 4$. We have

$$\begin{aligned} \frac{\partial h}{\partial x} &= \left(\frac{\partial h}{\partial S} \quad \frac{\partial h}{\partial I} \right) = \left(0 \quad \gamma \right), \\ \frac{\partial h}{\partial \Theta} &= \left(\frac{\partial h}{\partial \beta} \quad \frac{\partial h}{\partial \gamma} \quad \frac{\partial h}{\partial S_0} \quad \frac{\partial h}{\partial I_0} \right) = \left(0 \quad I \quad 0 \quad 0 \right), \end{aligned}$$

and

$$\frac{\partial x}{\partial \Theta} = \begin{pmatrix} \frac{\partial S}{\partial \beta} & \frac{\partial S}{\partial \gamma} & \frac{\partial S}{\partial S_0} & \frac{\partial S}{\partial I_0} \\ \frac{\partial I}{\partial \beta} & \frac{\partial I}{\partial \gamma} & \frac{\partial I}{\partial S_0} & \frac{\partial I}{\partial I_0} \end{pmatrix}.$$

This yields

$$\frac{\partial h}{\partial x} \frac{\partial x}{\partial \Theta} = \gamma \begin{pmatrix} \frac{\partial I}{\partial \beta} & \frac{\partial I}{\partial \gamma} & \frac{\partial I}{\partial S_0} & \frac{\partial I}{\partial I_0} \end{pmatrix}.$$

Therefore,

$$\chi = \gamma \begin{pmatrix} \frac{\partial I}{\partial \beta} & \frac{\partial I}{\partial \gamma} + I & \frac{\partial I}{\partial S_0} & \frac{\partial I}{\partial I_0} \end{pmatrix}.$$

Let $\{t_i\}$, $i = 0, 1, 2, \dots, M$, be the sampling times. Fisher's Information Matrix is defined as:

$$\text{FIM}(\Theta, \sigma) = \frac{1}{\sigma^2} \sum_{i=1}^M \chi(t_i, \Theta)^T \chi(t_i, \Theta),$$

where σ^2 is defined as in equation (5.4), since the initial conditions are unknown in this example.

Computing Computing Fisher's Information Matrix. We make the following decomposition:

$$\frac{\partial x}{\partial \Theta} = \begin{pmatrix} \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial x_0} \end{pmatrix} = \begin{pmatrix} z & w \end{pmatrix},$$

with

$$z = \frac{\partial x}{\partial \theta} = \begin{pmatrix} \frac{\partial S}{\partial \beta} & \frac{\partial S}{\partial \gamma} \\ \frac{\partial I}{\partial \beta} & \frac{\partial I}{\partial \gamma} \end{pmatrix}, \quad \text{and} \quad w = \frac{\partial x}{\partial x_0} = \begin{pmatrix} \frac{\partial S}{\partial S_0} & \frac{\partial S}{\partial I_0} \\ \frac{\partial I}{\partial S_0} & \frac{\partial I}{\partial I_0} \end{pmatrix}.$$

Letting

$$A(t) = \frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f_S}{\partial S} & \frac{\partial f_S}{\partial I} \\ \frac{\partial f_I}{\partial S} & \frac{\partial f_I}{\partial I} \end{pmatrix} = \begin{pmatrix} -\tilde{\beta}I & -\tilde{\beta}S \\ \tilde{\beta}I & \tilde{\beta}S - \gamma \end{pmatrix},$$

and

$$B(t) = \frac{\partial f}{\partial \theta} = \begin{pmatrix} \frac{\partial f_S}{\partial \beta} & \frac{\partial f_S}{\partial \gamma} \\ \frac{\partial f_I}{\partial \beta} & \frac{\partial f_I}{\partial \gamma} \end{pmatrix} = \begin{pmatrix} -SI & 0 \\ SI & -I \end{pmatrix}.$$

The FIM can be computed via numerically solving the following system of ODE's:

$$\begin{aligned} \dot{x} &= f(x, \theta), \quad x(0) = x_0, \\ \dot{z} &= Az + B, \quad z(0) = 0_{n \times p}, \\ \dot{w} &= Aw, \quad w(0) = \text{Id}_{n \times n}, \end{aligned}$$

which repeats equation (5.7). In the following code, the entries of x , z , and w are indexed in this way:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad z = \begin{pmatrix} z_3 & z_5 \\ z_4 & z_6 \end{pmatrix}, \quad w = \begin{pmatrix} w_7 & w_9 \\ w_8 & w_{10} \end{pmatrix},$$

leading to

$$\chi = \gamma \begin{pmatrix} z_4 & z_6 + x_2 & w_8 & w_{10} \end{pmatrix}.$$

D.2. Numerical implementation. Although the code is very similar the one provided in the previous example (Appendix C), we provide it for convenience, as it required a number of small changes.

We define the following functions in the Scilab environment:

```
function [kguess_n ,SSE]=identifKMK(OBS,T,kguess)
t0=T(1);m=length(T);
[x,SSE]=lsqrslv(kguess,errorKmcK,m,[1.d-8,1.d-8,1.d-5,1d9,0,100]);
nbparam=length(kguess);
kguess_n=x;
x0=[kguess_n(3);kguess_n(4)];
sol=ode(x0,t0,T,list(KmcK,x(1),x(2)));
kguess_n=kguess_n(:);
SSE=sum(SSE.^2);
xset("window",1);
sol1=ode(x0,t0,T(1):0.01:T($),list(KmcK ,x(1),x(2)));
clf;
plot((T(1):0.01:T($))',kguess_n(2)*sol1(2,:),'k');
plot(T',OBS ,'ko')
legend(["$\Large \gamma I(t)$", "$\Large \mbox{Data}$"])
ylabel("$\Large \mbox{number of deaths per week}$","fontsize",3);
xlabel("$\Large \mbox{time }t\mbox{ (in week)}$", "fontsize",3);
endfunction

function y=errorKmcK(k,m)
x0=[k(3);k(4)];
x0=x0(:);
B=k(1);
GAMMA=k(2);
sol=ode(x0,t0,T,list(KmcK,B,GAMMA));
predic=GAMMA*sol(2,:);
predic=predic(:);
OBS=OBS(:);
y=OBS-predic;
endfunction

function xdot=KmcK(t,x,B,GAMMA)
xdot=[-B*x(2),0; B*x(2),-GAMMA]*x
endfunction
```

Then, the Scilab session goes like this

```

--> load('databombay')
ans =
T
--> OBS=dataBSFlu

OBS =

column 1 to 9
8. 10. 12. 16. 24. 48. 51. 92. 124.

column 10 to 16
178. 280. 387. 442. 644. 779. 702.

column 17 to 23
695. 870. 925. 802. 578. 404. 296.

column 24 to 31
162. 106. 64. 46. 35. 27. 28. 24.

--> M=length(OBS);

--> T=0:M-1;

--> S0=15000;I0=7;gamma0=0.6;b0=8e-5;//Initial guesses

--> param=[b0,gamma0,S0,I0];//With b=beta/N

--> p=length(param);

--> [param,SSE]=identifKMK(OBS,T,param,N)

SSE =

106336.49

param =

0.0000855
3.7161743
48113.13
1.4213612

--> n=2;sigma2=SSE/(M-(n+p))

sigma2 =

```

```
4253.4597
```

```
--> B=param(1);GAMMA=param(2);
```

We then compute confidence intervals using the formulas (5.2), (5.3), (5.4), (5.5), (5.6).

```
function FIM=fimKmcK(x0,T,B,GAMMA,sigma2)
// Compute the sensitivity matrix
x0=x0(:);t0=T(1);
X0=[x0;zeros(4,1);eye(2,2)(:)];
sol=ode(X0,t0,T,list(JKmcK,B,GAMMA));
M=GAMMA*sol([4,6,8,10],:);
M(2,:)=M(2,)+GAMMA*sol(2,:);
FIM=M*M'./sigma2;
endfunction

function xdot=JKmcK(t,x,B,GAMMA)
xdot(1)=-B*x(1)*x(2);
xdot(2)=B*x(1)*x(2)-GAMMA*x(2);

xdot(3)=-B*x(2)*x(3)-B*x(1)*x(4)-x(1)*x(2);
xdot(4)=B*x(2)*x(3)+(B*x(1)-GAMMA)*x(4)+x(1)*x(2);
xdot(5)=-B*x(2)*x(5)-B*x(1)*x(6);
xdot(6)=B*x(2)*x(5)+(B*x(1)-GAMMA)*x(6)-x(2);

xdot(7)=-B*x(2)*x(7)-B*x(1)*x(8);
xdot(8)=B*x(2)*x(7)+(B*x(1)-GAMMA)*x(8);
xdot(9)=-B*x(2)*x(9)-B*x(1)*x(10);
xdot(10)=B*x(2)*x(9)+(B*x(1)-GAMMA)*x(10);
endfunction
```

Then the Scilab session is

```
--> x0=[param(3);param(4)];

--> FIM=fimKmcK(x0,T,B,GAMMA)
FIM =

    1.100D+14   -2.117D+09   196924.7   92108921.
   -2.117D+09   40885.251   -3.7835134   -1826.2845
    196924.7   -3.7835134   0.0003533   0.1603199
    92108921.   -1826.2845   0.1603199   118.22324

--> cond(FIM)
ans =

9.141D+24

--> CovMAT=inv(FIM)
Warning: Matrix is close to singular or badly scaled.
```

```

CovMAT =

    0.0000006    0.0093978   -220.96547   -0.0128142
    0.0093978    150.37071   -3535242.5   -204.997
   -220.96547   -3535242.5    8.313D+10    4820745.
   -0.0128142   -204.997    4820745.    279.63056

--> t=cdf("T",M-(n+p),0.975,0.025)
t =

2.0595386

--> confB=t*sqrt(CovMAT(1,1))
confB =

0.0015784

--> confGAMMA=t*sqrt(CovMAT(2,2))
confGAMMA =

25.255243

--> confS0=t*sqrt(CovMAT(3,3))
confS0 =

593794.26

--> confI0=t*sqrt(CovMAT(4,4))
confI0 =

34.439929

```

Acknowledgments. The authors are grateful to P.-A. Bliman, N. Cunniffe, C. Lobry, J. Harmand, T. Sari and M. Souza for exchanges and fruitful discussions.

REFERENCES

- [1] A. ABDELHEDI, D. BOUTAT, L. SBITA, R. TAMI, AND D.-Y. LIU, *Observer design for a class of nonlinear piecewise systems. Application to an epidemic model with treatment*, Math. Biosci., 271 (2016), pp. 128–135.
- [2] D. AEYELS, *Generic observability of differentiable systems*, SIAM J. Control Optim., 19 (1981), pp. 595–603.
- [3] ———, *On the number of samples necessary to achieve observability*, Systems Control Lett., 1 (1981/82), pp. 92–94.
- [4] H. AKAIKE, *A new look at the statistical model identification*, IEEE Trans. Autom. Control, 19 (1976), pp. 716–723.
- [5] S. ALONSO-QUESADA, M. DE LA SEN, R. P. AGARWAL, AND A. IBEAS, *An observer-based vaccination control law for an SEIR epidemic model based on feedback linearization techniques for nonlinear systems*, Adv. Difference Equ., (2012), pp. 2012:161, 32.
- [6] R. M. ANDERSON AND R. M. MAY, *Infectious Diseases of Humans. Dynamics and Control*, Oxford science publications, 1991.
- [7] R. M. ANDERSON, R. M. MAY, AND S. GUPTA, *Non-linear phenomena in host-parasite inter-*

- actions.*, Parasitology, 99 Suppl (1989), pp. S59–79.
- [8] B. ANDRÉA-NOVELL AND M. DE LARA, *Control Theory for Engineers: A Primer.*, Springer, Berlin, Heidelberg (2013).
- [9] V. ANDRIEU, G. BESANÇON AND U. SERRES, *Observability necessary conditions for the existence of observers.* IEEE Conference on Decision and Control, Florence, Italy (2013).
- [10] M. ANGUELOVA, *Observability and identifiability of nonlinear systems with applications in biology*, PhD thesis, Chalmers University of Technology and Göteborg University, 2007.
- [11] M. ANGUELOVA, J. KARLSSON, AND M. JIRSTRAND, *Minimal output sets for identifiability.*, Math Biosci, 239 (2012), pp. 139–153.
- [12] M. ASHYRALIYEV, Y. FOMEKONG-NANFACK, J. A. KAANDORP, AND J. G. BLOM, *Systems biology: parameter estimation for biochemical models.*, FEBS J, 276 (2009), pp. 886–902.
- [13] 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.
- [14] 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.
- [15] E. BALSACANTO AND J. R. BANGA, *Amigo, a toolbox for advanced model identification in systems biology using global optimization.*, Bioinformatics, 27 (2011), pp. 2311–2313.
- [16] E. BALSACANTO, D. HENRIQUES, A. GÁBOR, AND J. R. BANGA, *Amigo2, a toolbox for dynamic modeling, optimization and control in systems biology.*, Bioinformatics, 32 (2016), pp. 3357–3359.
- [17] 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.
- [18] 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.
- [19] H. T. BANKS, S. L. ERNSTBERGER, AND S. L. GROVE, *Standard errors and confidence intervals in inverse problems: sensitivity and associated pitfalls*, J. Inverse Ill-Posed Probl., 15 (2007), pp. 1–18.
- [20] 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.
- [21] H. T. BANKS AND H. T. TRAN, *Mathematical and experimental modeling of physical and biological processes*, Textbooks in Mathematics, CRC Press, Boca Raton, FL, 2009. With 1 CD-ROM (Windows, Macintosh and UNIX).
- [22] A. D. BARBOUR, *Macdonald’s model and the transmission of bilharzia.*, Trans R Soc Trop Med Hyg, 72 (1978), pp. 6–15.
- [23] R. BELLMAN, AND K.J. ÅSTRÖM, *On structural identifiability*, Math. Biosci., 7 (1970), pp. 329–339.
- [24] G. BELLU, M. P. 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.
- [25] D. BICHARA, N. COZIC, AND A. IGGIDR, *On the estimation of sequestered infected erythrocytes in plasmodium falciparum malaria patients*, Math. Biosci. Eng., 11 (2014), pp. 741–759.
- [26] N. A. BODUNOV, S. A. KOLBINA, AND S. Y. PLYUGIN, *Locally parameter identifiable systems are generic*, Vestnik St. Petersburg Univ. Math., 45 (2012), pp. 61–64.
- [27] B. M. BOLKER, *Ecological models and data in R*. Princeton University Press (2008).
- [28] ———, *Prevalence of locally parameter identifiable systems*, Vestnik St. Petersburg Univ. Math., 48 (2015), pp. 204–208.
- [29] F. BOULIER, F. OLLIVIER, D. LAZARD, AND M. PETITOT, *Computing representations for radicals of finitely generated differential ideals*, Appl. Algebra Engrg. Comm. Comput., 20 (2009), pp. 73–121.
- [30] 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.
- [31] 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.
- [32] F. BRAUER, J. WU, AND P. VAN DEN DRIESSCHE, eds., *Mathematical Epidemiology*, no. 1945 in Lectures Notes in Math., Springer-Verlag, 2008.
- [33] R. BRUN, P. REICHERT, AND KUNSCH, *Practical identifiability analysis of large environmental simulation models.*, Water Ressources Research, 37 (2001), pp. 1015–1030.
- [34] B. N. CANTÓ, C. COLL, AND E. SÁNCHEZ, *Estimation of parameters in a structured SIR*

- model*, Adv. Difference Equ., (2017), pp. Paper No. 33, 13.
- [35] 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.
- [36] M. A. CAPISTRAN, M. A. MORELES, AND B. LARA, *Parameter estimation of some epidemic models. The case of recurrent epidemics caused by respiratory syncytial virus*, Bull. Math. Biol., 71 (2009), pp. 1890–1901.
- [37] F. CELLE, J.-P. GAUTHIER, D. KAZAKOS, AND G. SALLET, *Synthesis of nonlinear observers: a harmonic-analysis approach*, Math. Systems Theory, 22 (1989), pp. 291–322.
- [38] C. D. S. CENTER, *Influenza in a boarding school*, BMJ, 1 (1978), p. 587.
- [39] W. CHEN AND M. SAIF, *Unknown input observer design for a class of nonlinear systems: an LMI approach*, American Control Conference, Minneapolis, MN, USA (2006).
- [40] O. CHİŞ, J. R. BANGA, AND E. BALSACANTO, *GenSSI: a software toolbox for structural identifiability analysis of biological models.*, Bioinformatics, 27 (2011), pp. 2610–2611.
- [41] 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.
- [42] O.-T. CHIS, A. F. VILLAVARDE, J. R. BANGA, AND E. BALSACANTO, *On the relationship between sloppiness and identifiability*, Mathematical biosciences, 282, 147–161, 2016.
- [43] O.-T. CHIS, A. F. VILLAVARDE, J. R. BANGA, AND E. BALSACANTO, *On the relationship between sloppiness and identifiability*, Math. Biosci., 282 (2016), pp. 147–161.
- [44] I.-C. CHOU AND E. O. VOIT, *Recent developments in parameter estimation and structure identification of biochemical and genomic systems*, Math. Biosci., 219 (2009), pp. 57–83.
- [45] 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.
- [46] 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.
- [47] 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.
- [48] C. COBELLI AND J. J. R. DI STEFANO, *Parameter and structural identifiability concepts and ambiguities: a critical review and analysis.*, Am J Physiol, 239 (1980), pp. R7–24.
- [49] C. COBELLI, A. LEPSCHY, AND G. ROMANIN-JACUR, *Identifiability of compartmental systems and related structural properties*, Math. Biosci., 44 (1979), pp. 1–18.
- [50] C. COBELLI AND G. ROMANIN-JACUR, *Controllability, Observability and Structural Identifiability of Multi Input and Multi Output Biological Compartmental Systems*, IEEE Trans Biomed Eng, 23 (1976), pp. 93–100.
- [51] ———, *On the structural identifiability of biological compartmental systems in a general input-output configuration*, Math. Biosci., 30 (1976), pp. 139–151.
- [52] D.J. COLE, *Parameter redundancy and identifiability*. Chapman and Hall/CRC, 2020.
- [53] COMMISSION, *Xxii-epidemiological observations made by the commission in bombay city*, J. Hyg (Lond), 7 (1907), pp. 724–798.
- [54] D. DALEY AND J. GANI, *Epidemic modelling : An introduction*, Cambridge University Press, 1999.
- [55] M. DAVIDIAN AND D. GILTINAN, *Nonlinear Models for Repeated Measurement Data.*, Chapman and Hall, Boca Raton, 1995.
- [56] 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.
- [57] M. DE LA SEN, A. IBEAS, AND S. ALONSO-QUESADA, *Observer-based vaccination strategy for a true mass action SEIR epidemic model with potential estimation of all the populations*, Discrete Dyn. Nat. Soc., (2011), pp. Art. ID 743067, 19.
- [58] P. DE LEENHEER AND H. L. SMITH, *Virus dynamics: A global analysis.*, SIAM J. Appl. Math., 63 (2003), pp. 1313–1327.
- [59] G. DE VRIES, T. HILLEN, M. LEWIS, J. MÜLLER, AND B. SCHÖNFISCH, *A course in mathematical biology*, vol. 12 of Mathematical Modeling and Computation, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006. Quantitative modeling with mathematical and computational methods.
- [60] J. W. DEMMEL, *On condition numbers and the distance to the nearest ill-posed problem*,

- Numer. Math., 51 (1987), pp. 251–289.
- [61] L. DENIS-VIDAL AND G. JOLY-BLANCHARD, *A necessary condition, a sufficient condition for structural identifiability*, in proceedings 36th CDC, 1997, pp. 1289–1290.
- [62] L. DENIS-VIDAL AND G. JOLY-BLANCHARD, *An easy to check criterion for (un)identifiability of uncontrolled systems and its applications*, IEEE Trans. Automat. Control, 45 (2000), pp. 768–771.
- [63] 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.
- [64] M. DIABY, A. IGGIDR, AND M. SY, *Observer design for a schistosomiasis model*, Math. Biosci., 269 (2015), pp. 17–29.
- [65] S. DIOP, *Elimination in control theory*, Math. Control Signals Systems, 4 (1991), pp. 17–32.
- [66] ———, *Differential-algebraic decision methods and some applications to system theory*, Theoret. Comput. Sci., 98 (1992), pp. 137–161. Second Workshop on Algebraic and Computer-theoretic Aspects of Formal Power Series (Paris, 1990).
- [67] S. DIOP AND M. FLIESS, *Nonlinear observability, identifiability, and persistent trajectories*, in proceedings 36th IEEE-CDC, 1991, pp. 714–719.
- [68] ———, *On nonlinear observability*, in Proceedings EEC91, vol. 1, Hermès, 1991, pp. 154–211.
- [69] J. DISTEFANO AND C. COBELLI, *On parameter and structural identifiability : nonunique observability/reconstructibility for identifiable systems, other ambiguities and new definitions*, IEEE Trans. Autom. Control, 25 (1980), pp. 830–833.
- [70] J. J. R. DISTEFANO, *Dynamic systems Biology modeling and application*, Academic Press, 2013.
- [71] S. DOBRE, T. BASTOGNE, C. PROFETA, M. BARBERI-HEYOB, AND A. RICHARD, *Limits of variance-based sensitivity analysis for non-identifiability testing in high dimensional dynamic models*, Automatica J. IFAC, 48 (2012), pp. 2740–2749.
- [72] A. DUCROT, P. MAGAL, T. NGUYEN, AND G. F. WEBB, *Identifying the number of unreported cases in SIR epidemic models*, Math. Med. Biol., 37 (2020), pp. 243–261.
- [73] 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.
- [74] N. D. EVANS, M. J. CHAPMAN, M. J. CHAPPELL, AND K. R. GODFREY, *Identifiability of uncontrolled nonlinear rational systems*, Automatica J. IFAC, 38 (2002), pp. 1799–1805.
- [75] N. D. EVANS, M. J. CHAPPELL, M. J. CHAPMAN, AND K. R. GODFREY, *Structural indistinguishability between uncontrolled (autonomous) nonlinear analytic systems*, Automatica J. IFAC, 40 (2004), pp. 1947–1953.
- [76] G. CICCARELLA, M. DALA MAARA AND A. GERMANI, *A Luenberger-like observer for nonlinear systems*, Inter. J. of Control, 57 (1993), pp. 537–556.
- [77] 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.
- [78] 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.
- [79] ———, *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.
- [80] ———, *Automatique et corps différentiels*, Forum Math., 1 (1989), pp. 227–238.
- [81] M. FLIESS AND T. GLAD, *Essays on Control*, no. 8 in Springe sciences, Springer, 1993, ch. An Algebraic Approach to Linear and Nonlinear Control, pp. 223–267.
- [82] M. FLIESS AND H. SIRA-RAMIREZ, *Reconstructeurs d’Etats*, CRAS, 338 (2004), pp. 91–96.
- [83] A. GÁBOR, A. F. VILLAVARDE, AND J. R. BANGA, *Parameter identifiability analysis and visualization in large-scale kinetic models of biosystems.*, BMC Syst Biol, 11 (2017), p. 54.
- [84] J. P. GAUTHIER AND I. A. K. KUPKA, *Observability for systems with more outputs than inputs and asymptotic observers*, Math. Z., 223 (1996), pp. 47–78.
- [85] J. P. GAUTHIER AND I. A. K. KUPKA, *Deterministic Observation Theory and Applications*, Cambridge University Press (2001).
- [86] G. H. GOLUB AND C. VAN LOAN, *Matrix Computations*, The John Hopkins Univesity Press, 1989.
- [87] M. B. GRAVENOR, A. L. LLOYD, P. G. KREMSNER, M. A. MISSINOU, M. ENGLISH, K. MARSH, AND D. KWIATKOWSKI, *A model for estimating total parasite load in falciparum malaria patients.*, J Theor Biol, 217 (2002), pp. 137–48.

- [88] M. B. GRAVENOR, A. R. MCLEAN, AND D. KWIATKOWSKI, *The regulation of malaria parasitaemia: parameter estimates for a population model.*, Parasitology, 110 (Pt 2) (1995), pp. 115–22.
- [89] M. B. GRAVENOR, M. B. VAN HENSBROEK, AND D. KWIATKOWSKI, *Estimating sequestered parasite population dynamics in cerebral malaria.*, Proc Natl Acad Sci U S A, 95 (1998), pp. 7620–4.
- [90] E. W. GRIFFITH AND K. S. P. KUMAR, *On the observability of nonlinear systems. I*, J. Math. Anal. Appl., 35 (1971), pp. 135–147.
- [91] J. GUILLAUME, J. JAKEMAN, S. MARSILI-LINELLI, M. ASHER, P. BRUNNER, B. CROKE, M. HILL, A. JAKEMAN, K. KEESAMN, S. RAZAVI, AND J. STIGTER, *Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose*, Environ. Model. Softw., 119 (2019), pp. 418–432.
- [92] K. P. HADELER, *Parameter identification in epidemic models*, Math. Biosci., 229 (2011), pp. 185–189.
- [93] 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.
- [94] R. HERMANN AND A. J. KRENER, *Nonlinear controllability and observability.*, IEEE Trans. Autom. Control, 22 (1977), pp. 728–740.
- [95] H. W. HETHCOTE AND J. W. VAN ARK, *Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation, and immunization programs*, Math. Biosci., 84 (1987), pp. 85–118.
- [96] 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.
- [97] H. HONG, A. OVCHINNIKOV, G. POGUDIN, AND C. YAP, *SIAN: a tool for assessing structural identifiability of parametric ODEs*, ACM Commun. Comput. Algebra, 53 (2019), pp. 37–40.
- [98] H. HONG, A. OVCHINNIKOV, G. POGUDIN, AND C. YAP, *Global identifiability of differential models*, Comm. Pure Appl. Math., 73 (2020), pp. 1831–1879.
- [99] M. HOU AND P. C. MÜLLER, *Design of observers for linear systems with unknown inputs*, IEEE Trans. Automat. Control, 37 (1992), pp. 871–875.
- [100] S. HUG, A. RAUE, J. HASENAUER, J. BACHMANN, U. KLINGMÜLLER, J. TIMMER, AND F. J. THEIS, *High-dimensional Bayesian parameter estimation: case study for a model of JAK2/STAT5 signaling*, Math. Biosci., 246 (2013), pp. 293–304.
- [101] A. IBEAS, M. DE LA SEN, S. ALONSO-QUESADA, AND I. ZAMANI, *Stability analysis and observer design for discrete-time SEIR epidemic models*, Adv. Difference Equ., (2015), pp. 2015:122, 21.
- [102] J. A. JACQUEZ, *Compartmental analysis in Biology and Medicine*, BioMedware, 1996.
- [103] 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.
- [104] D. L. I. JANZEN, L. BERGENHOLM, M. JIRSTRAND, J. PARKINSON, J. YATES, N. D. EVANS, AND M. J. CHAPPELL, *Parameter identifiability of fundamental pharmacodynamic models.*, Front Physiol, 7 (2016), p. 590.
- [105] 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).
- [106] M. JIRSTRAND, *Algebraic Methods for Modeling and Design in Control*, PhD thesis, Linkoping Studies in Science and Technology, Linkoping University, 1996.
- [107] G. JOLY-BLANCHARD AND L. DENIS-VIDAL, *Some remarks about an identifiability result of nonlinear systems*, Automatica, 34 (1998), pp. 1151–1152.
- [108] T. KAILATH, *Linear systems*; Prentice-Hall (1990).
- [109] R. E. KALMAN, *Mathematical description of linear dynamical systems*, J. SIAM Control Ser. A, 1 (1963), pp. 152–192 (1963).
- [110] 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.
- [111] J. T. KEMPER, *Identification of silent infections in SIR epidemics*, Bull. Math. Biol., 43 (1981), pp. 249–257.
- [112] W. KERMACK AND A. MCKENDRICK, *A contribution to the mathematical theory of epidemics*,

- Proc. R. Soc., A115 (1927), pp. 700–721.
- [113] E. R. KOLCHIN, *Differential algebra and algebraic groups*, Academic Press, New York-London, 1973. Pure and Applied Mathematics, Vol. 54.
- [114] T. C. KOOPMANS AND O. REIERSØ L, *The identification of structural characteristics*, Ann. Math. Statistics, 21 (1950), pp. 165–181.
- [115] F.-X. LE DIMET AND O. TALAGRAND, *Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects*, Tellus, 38A (1986), pp. 97–110.
- [116] J. LEANDER, T. LUNDH, AND M. JIRSTRAND, *Stochastic differential equations as a tool to regularize the parameter estimation problem for continuous time dynamical systems given discrete time measurements*, Math. Biosci., 251 (2014), pp. 54–62.
- [117] M. LI, J. DUSHOFF, AND B. M. BOLKER, *Fitting mechanistic epidemic models to data: a comparison of simple Markov chain Monte Carlo approaches*, Stat. Methods Med. Res., 27 (2018), pp. 1956–1967.
- [118] M. Y. LI, *An introduction to mathematical modeling of infectious diseases*, vol. 2 of Mathematics of Planet Earth, Springer, Cham, 2018.
- [119] 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.
- [120] L. LJUNG, *System identification : Theory for the user*, Prentice Hall, 1999.
- [121] L. LJUNG AND T. GLAD, *On global identifiability for arbitrary model parametrizations*, Automatica J. IFAC, 30 (1994), pp. 265–276.
- [122] D.G. LUENBERGER, *An Introduction to Observers*, IEEE Transactions on Automatic Control, 16, (1971), pp. 596–602.
- [123] G. MACDONALD, *The dynamics of helminth infections, with special reference to schistosomes.*, Trans R Soc Trop Med Hyg, 59 (1965), pp. 489–506.
- [124] P. MAGAL AND G. WEBB, *The parameter identification problem for SIR epidemic models: identifying unreported cases*, J. Math. Biol., 77 (2018), pp. 1629–1648.
- [125] G. MARGARIA, E. RICCOMAGNO, M. J. CHAPPELL, AND H. P. WYNN, *Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences*, Math. Biosci., 174 (2001), pp. 1–26.
- [126] G. MARGARIA, E. RICCOMAGNO, AND L. J. WHITE, *Structural identifiability analysis of some highly structured families of state space models using differential algebra*, J. Math. Biol., 49 (2004), pp. 433–454.
- [127] M. MARTCHEVA, *An introduction to mathematical epidemiology*, vol. 61 of Texts in Applied Mathematics, Springer, New York, 2015.
- [128] H. MCCALLUM, N. BARLOW, AND J. HONE, *How should pathogen transmission be modelled?*, Trends Ecol Evol, 16 (2001), pp. 295–300.
- [129] N. MESHKAT, C. ANDERSON, AND J. J. DiSTEFANO, III, *Alternative to Ritt’s pseudodivision for finding the input-output equations of multi-output models*, Math. Biosci., 239 (2012), pp. 117–123.
- [130] 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.
- [131] N. MESHKAT, C. E.-Z. KUO, AND J. R. DiSTEFANO, *On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and combos: a novel web implementation.*, PLoS One, 9 (2014), p. e110261.
- [132] 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.
- [133] N. C. MESHKAT, *A Differential Algebra Method for Eliminating Unidentifiability*, PhD thesis, UCLA, 2011. Thesis (Ph.D.)—University of California, Los Angeles.
- [134] 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.
- [135] J. MORENO AND D. DOCHAIN, *Global observability and detectability analysis of uncertain reaction systems*, International Journal of Control, 81(7) (2008), pp. 1062–1070.
- [136] T. G. MÜLLER, N. NOYKOVA, M. GYLLENBERG, AND J. TIMMER, *Parameter identification in dynamical models of anaerobic waste water treatment*, Math. Biosci., 177/178 (2002), pp. 147–160. Deterministic and stochastic modeling of biointeraction (West Lafayette, IN, 2000).
- [137] J. MURRAY, *Mathematical Biology I: An introduction*, vol. 17 of Interdisciplinary Applied Mathematics, Springer-Verlag, 2002.
- [138] M. NAKATSUI, A. SEDOGLAVIC, F. LEMAIRE, F. BOULIER, A. ÜRGÜPLÜ, AND K. HORIMOTO, *A general procedure for accurate parameter estimation in dynamic systems using new*

- estimation errors*, in Algebraic and numeric biology, vol. 6479 of Lecture Notes in Comput. Sci., Springer, Heidelberg, 2012, pp. 149–166.
- [139] S. NAZARI, *The Unknown Input Observer and its Advantages with Examples*, preprint arXiv:1504.07300 (2015).
- [140] 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.
- [141] M. A. NOWAK AND R. M. MAY, *virus dynamics. Mathematical principles of immunology and virology*, Oxford University Press, 2000.
- [142] F. OLLIVIER, *Le problème de l'identifiabilité structurelle globale : étude théorique, méthodes effectives et bornes de complexité.*, PhD thesis, Ecole Polytechnique, 1990.
- [143] F. OLLIVIER, *Standard bases of differential ideals*, in Applied algebra, algebraic algorithms and error-correcting codes (Tokyo, 1990), vol. 508 of Lecture Notes in Comput. Sci., Springer, Berlin, 1991, pp. 304–321.
- [144] P. D. O'NEILL, *A tutorial introduction to bayesian inference for stochastic epidemic models using markov chain monte carlo methods*, Mathematical Biosciences, 180 (2002), pp. 103–114.
- [145] A. OVCHINNIKOV, A. PILLAY, G. POGUDIN, AND T. SCANLON, *Multi-experiment parameter identifiability of odes and model theory*, 2020.
- [146] S. PANT, *Information sensitivity functions to assess parameter information gain and identifiability of dynamical systems.*, J R Soc Interface, 15 (2018).
- [147] S. PANT AND D. LOMBARDI, *An information-theoretic approach to assess practical identifiability of parametric dynamical systems*, Math. Biosci., 268 (2015), pp. 66–79.
- [148] M. PEIFER AND J. TIMMER, *Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting.*, IET Syst. Biol., (2007).
- [149] 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.
- [150] H. POHJANPALO, *System identifiability based on the power series expansion of the solution*, Math. Biosci., 41 (1978), pp. 21–33.
- [151] M. RAISSI, N. RAMEZANI, AND P. SESHAIYER, *On parameter estimation approaches for predicting disease transmission through optimization, deep learning and statistical inference methods*, Lett. Biomath., 6 (2019), p. 26.
- [152] A. RAPAPORT AND A. MALOUM, *Design of exponential observers for nonlinear systems by embedding*, International Journal of Robust and Nonlinear Control, 14 (2004), pp. 273–288.
- [153] A. RAUE, V. BECKER, U. KLINGMÜLLER, AND J. TIMMER, *Identifiability and observability analysis for experimental design in nonlinear dynamical models*, Chaos, 20 (2010), pp. 045105, 8.
- [154] A. RAUE, J. KARLSSON, M. P. SACCOMANI, M. JIRSTRAND, AND J. TIMMER, *Comparison of approaches for parameter identifiability analysis of biological systems.*, Bioinformatics, 30 (2014), pp. 1440–1448.
- [155] 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.
- [156] A. RAUE, C. KREUTZ, T. MAIWALD, U. KLINGMULLER, AND J. TIMMER, *Addressing parameter identifiability by model-based experimentation.*, IET Syst Biol, 5 (2011), pp. 120–130.
- [157] A. RAUE, B. STEIERT, M. SCHELKER, C. KREUTZ, T. MAIWALD, H. HASS, J. VANLIER, C. TÖNSING, L. ADLUNG, R. ENGESSER, W. MADER, T. HEINEMANN, J. HASENAUER, M. SCHILLING, T. HÖFER, E. KLIPP, F. THEIS, U. KLINGMÜLLER, B. SCHÖBERL, AND J. TIMMER, *Data2dynamics: a modeling environment tailored to parameter estimation in dynamical systems.*, Bioinformatics, 31 (2015), pp. 3558–3560.
- [158] J. REID, *Structural identifiability in linear-time invariant systems*, IEEE Trans. Autom. Control, 22 (1977), pp. 242–246.
- [159] J. F. RITT, *Differential algebra*, Dover Publications, Inc., New York, 1966.
- [160] M. G. ROBERTS, *An epidemic model with noisy parameters*, Math. Biosci., 287 (2017), pp. 36–41.
- [161] C. RODA, B. VARUGHESE, D. HAN, AND M. LI, *Why is it difficult to accurately predict the covid-19 epidemic?*, preprint, (2020).
- [162] W. C. RODA, *Bayesian inference for dynamical systems.*, Infect Dis Model, 5 (2020), pp. 221–232.
- [163] 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.
- [164] M. RODRIGUEZ-FERNANDEZ, J. A. EGEA, AND J. R. BANGA, *Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems.*, BMC Bioinformatics, 7 (2006), p. 483.
- [165] R. ROSS, *The prevention of malaria*, John Murray, 1911.
- [166] 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.
- [167] M. P. SACCOMANI, *An effective automatic procedure for testing parameter identifiability of HIV/AIDS models*, Bull. Math. Biol., 73 (2011), pp. 1734–1753.
- [168] M. P. SACCOMANI, *Structural vs practical identifiability in system biology*, in IWWBIO 2013, 2013, pp. 305–313.
- [169] 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.
- [170] 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.
- [171] A. SEDOGLAVIC, *A probabilistic algorithm to test local algebraic observability in polynomial time*, J. Symbolic Comput., 33 (2002), pp. 735–755. Computer algebra (London, ON, 2001).
- [172] A. SOLONEN, H. HAARIO, J. M. TCHUENCHE, AND H. RWEZURA, *Studying the identifiability of epidemiological models using MCMC*, Int. J. Biomath., 6 (2013), pp. 1350008, 18.
- [173] E. D. SONTAG, *On the length of inputs necessary in order to identify a deterministic linear system*, IEEE Trans. Automat. Control, 25 (1980), pp. 120–121.
- [174] E. D. SONTAG, *Mathematical control theory, deterministic finite dimensional systems*, no. 6 in Texts in Applied Mathematics, Springer-Verlag, 1990.
- [175] E. D. SONTAG, *Spaces of observables in nonlinear control*, in Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, 1995, pp. 1532–1545.
- [176] ———, *Critical points for least-squares problems involving certain analytic functions, with applications to sigmoidal nets*, Adv. Comput. Math., 5 (1996), pp. 245–268.
- [177] E. D. SONTAG, *For differential equations with r parameters, $2r + 1$ experiments are enough for identification*, J. Nonlinear Sci., 12 (2002), pp. 553–583.
- [178] E. D. SONTAG, *Dynamic compensation, parameter identifiability, and equivariances.*, PLoS Comput Biol, 13 (2017), p. e1005447.
- [179] 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.
- [180] R. C. SPEAR AND A. HUBBARD, *Parameter estimation and site-specific calibration of disease transmission models.*, Adv Exp Med Biol, 673 (2010), pp. 99–111.
- [181] O. TALAGRAND, *On the mathematics of data assimilation*, Tellus, 33 (1981), pp. 321–339.
- [182] R. TAMI, D. BOUTAT, G. ZHENG, AND F. KRATZ, *Parameters and state estimation for dengue epidemic model*, in 2014 European control conference (ECC^o, 2014, pp. 528–533.
- [183] A. TARANTOLA, *Inverse problem theory and methods for model parameter estimation*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.
- [184] K. THOMASETH AND C. COBELLI, *Generalized sensitivity functions in physiological system identification.*, Ann Biomed Eng, 27 (1999), pp. 607–616.
- [185] C. TÖNSING, J. TIMMER, AND C. KREUTZ, *Profile likelihood-based analyses of infectious disease models*, Stat. Methods Med. Res., 27 (2018), pp. 1979–1998.
- [186] E. T. TUNALI AND T. J. TARN, *New results for identifiability of nonlinear systems*, IEEE Trans. Automat. Control, 32 (1987), pp. 146–154.
- [187] N. TUNCER, H. GULBUDAK, V. L. CANNATARO, AND M. MARCHEMA, *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.
- [188] N. TUNCER AND T. T. LE, *Structural and practical identifiability analysis of outbreak models*, Math. Biosci., 299 (2018), pp. 1–18.
- [189] N. TUNCER, M. MARCHEMA, B. LABARRE, AND S. PAYOUTE, *Structural and practical identifiability analysis of Zika epidemiological models*, Bull. Math. Biol., 80 (2018), pp. 2209–2241.
- [190] S. VAJDA, K. R. GODFREY, AND H. RABITZ, *Similarity transformation approach to identifiability analysis of nonlinear compartmental models*, Math. Biosci., 93 (1989), pp. 217–248.
- [191] J. X. VELASCO-HERNÁNDEZ, J. A. GARCÍA, AND D. KIRSCHNER, *Remarks on modeling host-*

- viral dynamics and treatment.*, in Mathematical approaches for emerging and reemerging infectious diseases: An introduction. Proceedings of a tutorial Introduction to epidemiology and immunology, C. e. e. a. Castillo-Chavez, ed., Springer. IMA Vol. Math. Appl. 125, 2002, pp. 287–308.
- [192] A. VILLAVERDE, *Observability and structural identifiability of nonlinear biological systems*, Complexity, 2019 (2019).
- [193] A. VILLAVERDE AND J. BANGA, *Reverse engineering and identification insystems biology: strategies, perspectivesand challenges*, J. R. Soc. Interface, 11 (2014).
- [194] A. F. VILLAVERDE, A. BARREIRO, AND A. PAPACHRISTODOULOU, *Structural identifiability of dynamic systems biology models.*, PLoS Comput Biol, 12 (2016), p. e1005153.
- [195] A. F. VILLAVERDE, N. TSIANTIS, AND J. R. BANGA, *Full observability and estimation of unknown inputs, states and parameters of nonlinear biological models.*, J R Soc Interface, 16 (2019), p. 20190043.
- [196] E. WALTER, *Identifiability of state space models*, vol. 46 of Lecture Notes in Biomathematics, Springer-Verlag, Berlin-New York, 1982. With applications to transformation systems.
- [197] 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.
- [198] E. WALTER AND L. PRONZATO, *Identification de Modèles Paramétriques à partir de Données Expérimentales*, Masson, 1994.
- [199] E. WALTER AND L. PRONZATO, *Identifiabilities and nonlinearities*, in Nonlinear systems, Vol. 1, Chapman and Hall, London, 1995, pp. 111–143.
- [200] 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.
- [201] Y. WANG AND E. D. SONTAG, *On two definitions of observation spaces.*, Syst. Control Lett., 13 (1989), pp. 279–289.
- [202] Y. WANG AND E. D. SONTAG, *Orders of input/output differential equations and state-space dimensions*, SIAM J. Control Optim., 33 (1995), pp. 1102–1126.
- [203] F.-G. WIELAND, A. L. HAUBER, M. ROSENBLATT, C. TÖNSING, AND J. TIMMER, *On structural and practical identifiability*, Curr Opin Syst Biol, 25 (2021), pp. 60–69.
- [204] 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.
- [205] X. XIA, *Estimation of HIV/AIDS parameters*, Automatica J. IFAC, 39 (2003), pp. 1983–1988.
- [206] 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.
- [207] J. W. T. YATES, R. D. O. JONES, M. WALKER, AND S. Y. A. CHEUNG, *Structural identifiability and indistinguishability of compartmental models.*, Expert Opin Drug Metab Toxicol, 5 (2009), pp. 295–302.