15th Triennial World Congress, Barcelona, Spain

THE STRUCTURAL IDENTIFIABILITY OF A GENERAL EPIDEMIC (SIR) MODEL WITH SEASONAL FORCING

N. D. Evans * **M. J. Chapman** ** **M. J. Chappell** * **K. R. Godfrey** *

* *School of Engineering, University of Warwick, COVENTRY, CV4 7AL UK*
** *School of MIS-Mathematics, Coventry University, COVENTRY, CV1 5FB UK*

Abstract: In this paper it is shown that a general SIR epidemic model, with the force of infection subject to seasonal variation, and a proportion of the number of infectives measured, is unidentifiable. This means that an uncountable number of different parameter vectors can, theoretically, give rise to the same idealised output data. Any subsequent parameter estimation from real data must be viewed with less confidence as a result. The approach is essentially that developed by Evans *et al.* (2002), with modifications to allow for time-variation in the effective contact rate. This approach utilises the existence of an infinitely differentiable transformation that connects the state trajectories corresponding to parameter vectors that give rise to identical output data.

Keywords: Structural identifiability, nonlinear systems, SIR models, seasonal variation.

1. INTRODUCTION

When developing a parametric state space model of a given physical system, such as the spread of an infectious disease, the important question of whether the parameters can be determined uniquely from experimental data arises. Structural identifiability is concerned with the answer to this question in the theoretical situation of noise-free, perfect and continuous measurement/observation data (see, for example, Bellman and Åström (1970) and the papers in Walter (1987)). This is a problem relating to the structure of the mathematical model, and deals with whether different vectors of parameters give rise to the same (ideal) measurement data. As such, structural identifiability is an important theoretical prerequisite to experiment design and parameter estimation (in which the aim is to determine the parameters from real, and possibly noisy, data).

The identifiability of linear models is a well understood concept, with many analytical techniques available (see Walter (1987), and in particular, the paper therein by Godfrey and DiStefano III (1987)). For nonlinear systems, there are relatively few techniques (the Taylor series approach (Pohjanpalo, 1978), the similarity transformation approach (Tunali and Tarn, 1987; Vajda *et al.*, 1989), and differential algebra techniques (Fliess and Glad, 1993; Ljung and Glad, 1994) being among the most common) and significant problems still remain to be overcome. One such problem results from the practical consideration that unique (global) identifiability be sought with respect to a specific input (Chappell *et al.*, 1990), for example a bolus (impulsive) input of drug in pharmacokinetics.

A large class of models falling into this category is that consisting of those without inputs (i.e., an identically zero input), for example general SIR (susceptible, infected, removed) type models used in the study of epidemics (Capasso, 1993). The identifiability of models of this form is typically approached via a Taylor series expansion of the output (Pohjanpalo, 1978). However, the paper by Evans *et al.* (2002) proposes a new method which utilises the existence of a smooth map that connects the state trajectories of different parameterisations that yield identical outputs. In this paper it is shown how this general theory can be applied to a general SIR model with seasonal variation in the force of infection. This general SIR model, together with the modifications necessary to allow for the temporal variation, is described in the next section.

2. GENERAL SIR MODEL

The general SIR model, which was introduced by Kermack and McKendrick (1927), is intended to describe the evolution of an epidemic in a constant and isolated host population of size N. The population is divided into three distinct classes: The susceptible class, with population size S, consisting of those hosts that can contract the disease and become infectious; the infectious class, with population size I, consisting of those hosts that have been infected and can transmit the disease to other hosts; and the removed class, with population R, consisting of those hosts that have contracted the disease and then died, or recovered with full immunity against subsequent reinfection.

The 'law of mass action' is assumed for the infection process ($S \longrightarrow I$) so that the *per capita* rate at which susceptibles are infected is proportional to the number of infectives. The constant of proportionality, denoted by β, is a crucial parameter in the model, called the effective contact rate. To preserve the constant population size it is assumed that the birth rate (into the susceptible class) matches the net mortality rate (μN). The SIR model is then given by the following system of ordinary differential equations (Capasso, 1993):

$$\dot{S} = \mu N - S\left(\mu + \beta I\right), \tag{1}$$

$$\dot{I} = I\left(\beta S - (\mu + \sigma)\right), \tag{2}$$

$$\dot{R} = \sigma I - \mu R \tag{3}$$

where $N = S(t) + I(t) + R(t)$ is a positive constant, $1/\mu$ is the average life expectancy, $1/\sigma$ is the mean infectious period, and β represents the average fraction of susceptibles contacted by a single infective that then contract the disease. It is not really necessary to include an equation for R since it does not appear in either of the first two equations, except through the sum N. Moreover, the population sizes S and I can be scaled to give the proportions, $s = S/N$ and $q = I/N$, of susceptibles and infectives respectively. The parameters μ, β and σ, together with the initial conditions and the constant N, are all assumed to be positive and unknown.

Each of the parameters in the model has practical significance and so it is essential to establish identifiability before any parameter fitting can be performed with confidence. Moreover, the parameter combination $R_0 = \beta N/(\mu + \sigma)$, called the *basic reproduction rate*, is the average number of secondary cases resulting from one primary infection when introduced into a completely susceptible population and is crucial when considering control of the disease: If $R_0 < 1$ the number of newly infected cases decreases to zero and the disease dies out, and if $R_0 > 1$ an epidemic occurs. Thus the aim of any control strategy is to reduce R_0 below 1. Typically this might be achieved by blocking the transmission of the disease, i.e., reducing β, through vaccination and/or culling. Hence it is essential that β be identifiable in order to design an effective strategy.

The identifiability of the model (1)–(3) will be analysed for the case where a proportion of the number of infectives is measured. This output is appropriate when the ratio of the number of cases to the number of infectious individuals in the population is assumed to be constant, and the data consist of the reported number of cases. For example, this was the situation assumed in Weber *et al.* (2001) for respiratory syncytial virus. The corresponding output for the model (1)–(3) is given by:

$$y = kI \tag{4}$$

where the parameter k is positive and unknown.

A further level of complication is added in models for diseases that have a force of infection that is subject to seasonal, or other temporal, fluctuations. For these models the constant effective contact rate, β, is replaced by a periodic function of time, $\beta(t)$ say. For example, Grenfell *et al.* (1995) modelled the dynamics of measles using the function

$$\beta(t) = \beta_0\left(1 + \beta_1 \cos\left(2\pi t\right)\right). \tag{5}$$

With t measured in years this corresponds to a function with an annual period. This will be the function used in this paper for the effective contact rate. To ensure that $\beta(t) > 0$, for all $t \geq 0$, it is assumed that $|\beta_1| < 1$.

3. STRUCTURAL IDENTIFIABILITY

In this section, nonlinear systems of the following form are considered:

$$\dot{\boldsymbol{x}}(t, \boldsymbol{p}) = \boldsymbol{f}(\boldsymbol{x}(t, \boldsymbol{p}), \boldsymbol{p}), \tag{6}$$

$$\boldsymbol{x}(0, \boldsymbol{p}) = \boldsymbol{x}_0(\boldsymbol{p}), \tag{7}$$

$$\boldsymbol{y}(t, \boldsymbol{p}) = \boldsymbol{h}(\boldsymbol{x}(t, \boldsymbol{p}), \boldsymbol{p}), \tag{8}$$

where $\boldsymbol{p} \in \Omega$, an open subset of $\mathbb{R}^q$, is a constant parameter vector, and the output $\boldsymbol{y}(t, \boldsymbol{p}) \in \mathbb{R}^r$. For all $\boldsymbol{p} \in \Omega$, denote by $M(\boldsymbol{p})$ the largest connected open subset of $\mathbb{R}^n$ containing $\boldsymbol{x}_0(\boldsymbol{p})$ such that both $\boldsymbol{f}(\cdot, \boldsymbol{p})$ and $\boldsymbol{h}(\cdot, \boldsymbol{p})$ are analytic on $M(\boldsymbol{p})$. Let $\tau(\boldsymbol{p})$ be the supremum of the set of all $\tau > 0$ such that $\boldsymbol{x}(t, \boldsymbol{p}) \in M(\boldsymbol{p})$ for $0 \leq t \leq \tau$.

Definition 1. Given a parameter vector $\boldsymbol{p}$, then $\overline{\boldsymbol{p}} \in \Omega$ is said to be *indistinguishable* from $\boldsymbol{p}$, written $\overline{\boldsymbol{p}} \sim \boldsymbol{p}$, if $\boldsymbol{y}(t, \boldsymbol{p}) = \boldsymbol{y}(t, \overline{\boldsymbol{p}})$ for all $t \in [0, \tau(\boldsymbol{p}))$. (In particular, this means that $\tau(\overline{\boldsymbol{p}}) \geq \tau(\boldsymbol{p})$.)

If $\overline{\boldsymbol{p}} \sim \boldsymbol{p}$, for $\boldsymbol{p}, \overline{\boldsymbol{p}} \in \Omega$, then it is not possible to distinguish between the parameter vectors $\boldsymbol{p}$ and $\overline{\boldsymbol{p}}$

from ideal output data on $[0, \tau(\boldsymbol{p}))$. With respect to this definition, the identifiability problem is to determine whether there exist other parameter vectors that are indistinguishable from a given one.

Definition 2. A model of the form (6)–(8) is said to be *globally identifiable at* $\boldsymbol{p} \in \Omega$ if $\overline{\boldsymbol{p}} \in \Omega$ and $\overline{\boldsymbol{p}} \sim \boldsymbol{p}$ imply that $\boldsymbol{p} = \overline{\boldsymbol{p}}$. If this is true on some suitably small neighbourhood of $\boldsymbol{p}$ then the model is *locally identifiable at* $\boldsymbol{p} \in \Omega$.

If the system (6)–(8) is globally identifiable at $\boldsymbol{p}$, then the resulting (ideal) output data from the model are unique.

Definition 3. If (6)–(8) is globally (locally) identifiable at $\boldsymbol{p}$ for all $\boldsymbol{p} \in \Omega$, except for a subset of a closed set of (Lebesgue) measure zero, then it is said to be *structurally globally (locally) identifiable*. The model is said to be *unidentifiable* if it is not structurally locally identifiable.

The *Lie derivative* of $h \in C^\infty(M(\boldsymbol{p}))$ (i.e., h is real-valued and infinitely differentiable on $M(\boldsymbol{p})$) along the vector field $\boldsymbol{f}$ is the smooth function given by

$$\mathrm{L}_{\boldsymbol{f}} h(\boldsymbol{x}) = \frac{\partial h}{\partial \boldsymbol{x}}(\boldsymbol{x}) \boldsymbol{f}(\boldsymbol{x}).$$

Let $\boldsymbol{f}^{\boldsymbol{p}}(\cdot) = \boldsymbol{f}(\cdot, \boldsymbol{p})$ and, for l, $1 \leq l \leq r$, $h_l^{\boldsymbol{p}}(\cdot) = h_l(\cdot, \boldsymbol{p})$. The system (6)–(8) is said to satisfy the Observability Rank Condition (ORC) at $\boldsymbol{x}_0(\boldsymbol{p})$ if there exist functions $\mu_1(\boldsymbol{x}, \boldsymbol{p}), \ldots, \mu_n(\boldsymbol{x}, \boldsymbol{p})$ of the form $h_j(\boldsymbol{x}, \boldsymbol{p})$, for some j, or $\mathrm{L}_{\boldsymbol{f}^{\boldsymbol{p}}}^m h_l^{\boldsymbol{p}}(\boldsymbol{x})$, for some l and m, such that the Jacobian matrix, evaluated at $\boldsymbol{x}_0(\boldsymbol{p})$, of the function defined by

$$H_{\boldsymbol{p}} : \boldsymbol{x} \longmapsto (\mu_1(\boldsymbol{x}, \boldsymbol{p}), \ldots, \mu_n(\boldsymbol{x}, \boldsymbol{p}))^\top$$

is nonsingular (see the paper by Hermann and Krener (1977) for more details, particularly in the context of observability).

If a system satisfies the ORC at the initial condition, for a particular parameter vector $\boldsymbol{p}$, then it is possible to construct a smooth mapping from the state corresponding to a parameter vector indistinguishable from $\boldsymbol{p}$ to the state corresponding to $\boldsymbol{p}$. The following result then lays the foundation for a method for testing the identifiability of system (6)–(8).

Theorem 4. Suppose that system (6)–(8) satisfies the ORC at $\boldsymbol{x}_0(\boldsymbol{p})$ for some $\boldsymbol{p} \in \Omega$. If $\overline{\boldsymbol{p}} \in \Omega$ is such that $\overline{\boldsymbol{p}} \sim \boldsymbol{p}$, then there exists an open neighbourhood $V_{\overline{\boldsymbol{p}}}$ of $\boldsymbol{x}_0(\overline{\boldsymbol{p}})$ and a smooth map $\lambda : V_{\overline{\boldsymbol{p}}} \longrightarrow \lambda(V_{\overline{\boldsymbol{p}}})$ such that

$$H_{\boldsymbol{p}}(\lambda(\boldsymbol{x})) = H_{\overline{\boldsymbol{p}}}(\boldsymbol{x}), \tag{9}$$

for all $\boldsymbol{x} \in V_{\overline{\boldsymbol{p}}}$, and

$$\lambda(\boldsymbol{x}_0(\overline{\boldsymbol{p}})) = \boldsymbol{x}_0(\boldsymbol{p}), \tag{10}$$

$$\boldsymbol{f}^{\boldsymbol{p}}(\lambda(\boldsymbol{x}(t, \overline{\boldsymbol{p}}))) = \frac{\partial \lambda}{\partial \boldsymbol{x}}(\boldsymbol{x}(t, \overline{\boldsymbol{p}})) \boldsymbol{f}^{\overline{\boldsymbol{p}}}(\boldsymbol{x}(t, \overline{\boldsymbol{p}})), \tag{11}$$

$$\boldsymbol{h}^{\boldsymbol{p}}(\lambda(\boldsymbol{x}(t, \overline{\boldsymbol{p}}))) = \boldsymbol{h}^{\overline{\boldsymbol{p}}}(\boldsymbol{x}(t, \overline{\boldsymbol{p}})), \tag{12}$$

for all $t \in [0, \tau(\boldsymbol{p}))$ with $\boldsymbol{x}(t, \overline{\boldsymbol{p}}) \in V_{\overline{\boldsymbol{p}}}$, where $\boldsymbol{x}(t, \overline{\boldsymbol{p}})$ is the solution of system (6)–(8) for $\overline{\boldsymbol{p}}$.

PROOF. Since (6)–(8) satisfies the ORC at $\boldsymbol{x}_0(\boldsymbol{p})$ there exists an open neighbourhood W of $\boldsymbol{x}_0(\boldsymbol{p})$ such that $H_{\boldsymbol{p}}$ is a diffeomorphism from W to $H_{\boldsymbol{p}}(W)$.

Since $\overline{\boldsymbol{p}} \sim \boldsymbol{p}$, it is seen that $(1 \leq j \leq r)$

$$h_j(\boldsymbol{x}(t, \boldsymbol{p})) = y_j(t, \boldsymbol{p}) = y_j(t, \overline{\boldsymbol{p}}) = h_j(\boldsymbol{x}(t, \overline{\boldsymbol{p}}))$$

for all $t \in [0, \tau(\boldsymbol{p}))$, and so

$$H_{\boldsymbol{p}}(\boldsymbol{x}(t, \boldsymbol{p})) = H_{\overline{\boldsymbol{p}}}(\boldsymbol{x}(t, \overline{\boldsymbol{p}})) \tag{13}$$

for all $t \in [0, \tau(\boldsymbol{p}))$. Setting $t = 0$ in this equation shows that $H_{\overline{\boldsymbol{p}}}(\boldsymbol{x}_0(\overline{\boldsymbol{p}})) \in H_{\boldsymbol{p}}(W)$. Therefore there exists a neighbourhood $V_{\overline{\boldsymbol{p}}}$ of $\boldsymbol{x}_0(\overline{\boldsymbol{p}})$ such that $H_{\overline{\boldsymbol{p}}}(V_{\overline{\boldsymbol{p}}}) \subset H_{\boldsymbol{p}}(W)$ and a smooth map $\lambda : V_{\overline{\boldsymbol{p}}} \longrightarrow \lambda(V_{\overline{\boldsymbol{p}}})$ defined by

$$\lambda(\boldsymbol{x}) = H_{\boldsymbol{p}}^{-1}(H_{\overline{\boldsymbol{p}}}(\boldsymbol{x})).$$

It is seen from its construction and (13), that this map satisfies (9), (10) and (12) (since $\boldsymbol{h}(\boldsymbol{x}(t, \boldsymbol{p})) = \boldsymbol{h}(\boldsymbol{x}(t, \overline{\boldsymbol{p}}))$ for $0 \leq t < \tau(\boldsymbol{p})$). For equation (11) note that

$$\begin{aligned} \boldsymbol{f}^{\boldsymbol{p}}(\lambda(\boldsymbol{x}(t, \overline{\boldsymbol{p}}))) = \dot{\boldsymbol{x}}(t, \boldsymbol{p}) &= \frac{\partial \lambda}{\partial \boldsymbol{x}}(\boldsymbol{x}(t, \overline{\boldsymbol{p}})) \dot{\boldsymbol{x}}(t, \overline{\boldsymbol{p}}) \\ &= \frac{\partial \lambda}{\partial \boldsymbol{x}}(\boldsymbol{x}(t, \overline{\boldsymbol{p}})) \boldsymbol{f}^{\overline{\boldsymbol{p}}}(\boldsymbol{x}(t, \overline{\boldsymbol{p}})) \end{aligned}$$

for all $t \in [0, \tau(\boldsymbol{p}))$ with $\boldsymbol{x}(t, \overline{\boldsymbol{p}}) \in V_{\overline{\boldsymbol{p}}}$. □

Suppose that the system (6)–(8) satisfies the ORC at $\boldsymbol{x}_0(\boldsymbol{p})$, for some $\boldsymbol{p} \in \Omega$. Denote by $\mathcal{S}(\boldsymbol{p})$ the subset of Ω, containing $\boldsymbol{p}$, of all possible parameter vectors $\overline{\boldsymbol{p}}$ such that λ, defined in (9), satisfies (10)–(12). If this set consists only of $\boldsymbol{p}$ then the system (6)–(8) is globally identifiable at $\boldsymbol{p}$. However, if $\mathcal{S}(\boldsymbol{p})$ consists of more vectors than just $\boldsymbol{p}$, it cannot be concluded that the system is not globally identifiable.

Remark 5. Let W be any open subset, containing $\boldsymbol{x}_0(\boldsymbol{p})$, of $M(\boldsymbol{p})$ such that $H_{\boldsymbol{p}}$ is a diffeomorphism when restricted to W. If $\boldsymbol{x}(t, \boldsymbol{p}) \in W$ for all $t \in [0, \tau(\boldsymbol{p}))$ and $\overline{\boldsymbol{p}} \sim \boldsymbol{p}$ (some $\overline{\boldsymbol{p}} \in \Omega$), then the neighbourhood $V_{\overline{\boldsymbol{p}}}$ of $\boldsymbol{x}_0(\overline{\boldsymbol{p}})$ (in Theorem 4) can be chosen such that $\boldsymbol{x}(t, \overline{\boldsymbol{p}}) \in V_{\overline{\boldsymbol{p}}}$ for all $t \in [0, \tau(\boldsymbol{p}))$.

This remark gives rise to the following corollary to Theorem 4.

Corollary 6. Suppose that system (6)–(8) satisfies the ORC at $\boldsymbol{x}_0(\boldsymbol{p})$ for some $\boldsymbol{p} \in \Omega$, and that there exists an open set W such that $\boldsymbol{x}(t, \boldsymbol{p}) \in W$, for all $t \in [0, \tau(\boldsymbol{p}))$, and $H_{\boldsymbol{p}} : W \longrightarrow H_{\boldsymbol{p}}(W)$ is a diffeomorphism. Then $\overline{\boldsymbol{p}} \sim \boldsymbol{p}$, $\overline{\boldsymbol{p}} \in \Omega$, if and only if there exists an open neighbourhood $V_{\overline{\boldsymbol{p}}}$ of $\boldsymbol{x}_0(\overline{\boldsymbol{p}})$ and a smooth map $\lambda : V_{\overline{\boldsymbol{p}}} \longrightarrow \lambda(V_{\overline{\boldsymbol{p}}})$ that satisfies equation (9) for all $\boldsymbol{x} \in V_{\overline{\boldsymbol{p}}}$, and equations (10)–(12) for all $t \in [0, \tau(\boldsymbol{p}))$.

PROOF. Suppose that $\overline{\boldsymbol{p}} \sim \boldsymbol{p}$ for some $\overline{\boldsymbol{p}} \in \Omega$. From Remark 5 it is seen that $V_{\overline{\boldsymbol{p}}}$ can be chosen such that $\boldsymbol{x}(t, \overline{\boldsymbol{p}}) \in V_{\overline{\boldsymbol{p}}}$ for all $t \in [0, \tau(\boldsymbol{p}))$. Hence applying Theorem 4 it is seen that $\lambda = H_{\boldsymbol{p}}^{-1} \circ H_{\overline{\boldsymbol{p}}}$ satisfies (9)–(12) for all $t \in [0, \tau(\boldsymbol{p}))$.

Conversely, let $V_{\overline{\boldsymbol{p}}}$ be a neighbourhood of $\boldsymbol{x}_0(\overline{\boldsymbol{p}})$ and $\lambda : V_{\overline{\boldsymbol{p}}} \longrightarrow \lambda\left(V_{\overline{\boldsymbol{p}}}\right)$ be a smooth map that satisfies (9), and (10)–(12) for all $t \in [0, \tau(\boldsymbol{p}))$. Defining $\boldsymbol{z}(t) = \lambda(\boldsymbol{x}(t, \overline{\boldsymbol{p}}))$ it is seen that $\boldsymbol{z}(\cdot)$ is the solution of the following system:

$$\dot{\boldsymbol{z}}(t) = \frac{\partial \lambda}{\partial \boldsymbol{x}}(\boldsymbol{x}(t, \overline{\boldsymbol{p}}))\boldsymbol{f}^{\overline{\boldsymbol{p}}}(\boldsymbol{x}(t, \overline{\boldsymbol{p}})) = \boldsymbol{f}^{\boldsymbol{p}}(\boldsymbol{z}(t)),$$
$$\boldsymbol{z}(0) = \lambda(\boldsymbol{x}_0(\overline{\boldsymbol{p}})) = \boldsymbol{x}_0(\boldsymbol{p})$$

and so, by the uniqueness of solutions, $\boldsymbol{z}(t) = \boldsymbol{x}(t, \boldsymbol{p})$. Moreover, equation (12) implies that

$$\boldsymbol{y}(t, \overline{\boldsymbol{p}}) = \boldsymbol{y}(t, \boldsymbol{p})$$

for all $t \in [0, \tau(\boldsymbol{p}))$. Hence $\overline{\boldsymbol{p}} \sim \boldsymbol{p}$. □

In the case of Corollary 6, the set $\mathcal{S}(\boldsymbol{p})$ consists of precisely those parameter vectors that are indistinguishable from $\boldsymbol{p}$. If there exists a neighbourhood, N, of $\boldsymbol{p}$ such that $\mathcal{S}(\boldsymbol{p}) \cap N = \{\boldsymbol{p}\}$ then the system is locally identifiable at $\boldsymbol{p}$. These results are structural if they remain true for all parameter vectors $\boldsymbol{p}$ except possibly where the components of $\boldsymbol{p}$ satisfy some *a priori* algebraic equation.

4. IDENTIFIABILITY ANALYSIS

Before considering the case with seasonal forcing an identifiability analysis for a general SIR model, with a proportion of the infectives measured, is presented. The model equations, (1) and (2) (since (3) is not necessary), together with the output (4), are of the form (6)–(8) and so Theorem 4, or Corollary 6, can be used in the structural identifiability analysis.

Example 7. The general SIR model (1)–(2), together with the output (4), can be written in the form

$$\dot{x}_1 = p_1 p_4 - x_1\left(p_1 + p_2 x_2\right) \qquad x_1(0) = p_5$$
$$\dot{x}_2 = x_2\left(p_2 x_1 - (p_1 + p_3)\right) \qquad x_2(0) = p_6$$
$$y = p_7 x_2$$

where $x_1 = S$, $x_2 = I$, $p_1 = \mu$, $p_2 = \beta$, $p_3 = \sigma$, $p_4 = N$, and $p_7 = k$. Let the vector of unknown parameters be denoted by $\boldsymbol{p} = (p_1, \ldots, p_7)^\top$ and the set of possible $\boldsymbol{p}$ be given by $\Omega = \left\{\boldsymbol{p} \in \mathbb{R}^7 : p_i > 0\right\}$. It is seen that $M(\boldsymbol{p}) = \mathbb{R}^2$.

To see that the ORC is satisfied at $\boldsymbol{x}_0(\boldsymbol{p})$ let $\mu_1(\boldsymbol{x}, \boldsymbol{p}) = p_7 x_2$, and

$$\mu_2(\boldsymbol{x}, \boldsymbol{p}) = \mathrm{L}_{\boldsymbol{f}^{\boldsymbol{p}}} \mu_1^{\boldsymbol{p}}(\boldsymbol{x}) = p_7 x_2\left(p_2 x_1 - (p_1 + p_3)\right),$$

where $\mu_1^{\boldsymbol{p}}(\boldsymbol{x}) = \mu_1(\boldsymbol{x}, \boldsymbol{p})$. The Jacobian matrix of the function defined by $H_{\boldsymbol{p}}(\boldsymbol{x}) = (\mu_1(\boldsymbol{x}, \boldsymbol{p}), \mu_2(\boldsymbol{x}, \boldsymbol{p}))^\top$ is given by

$$\frac{\partial H_{\boldsymbol{p}}}{\partial \boldsymbol{x}}(\boldsymbol{x}) = \begin{pmatrix} 0 & p_7 \\ p_2 p_7 x_2 & p_7(p_2 x_1 - (p_1 + p_3)) \end{pmatrix}.$$

Given any $\boldsymbol{p} \in \Omega$ this matrix has full rank for all $\boldsymbol{x} \in W = \{\boldsymbol{x} \in \mathbb{R}^2 : x_2 \neq 0\}$. Since $\boldsymbol{x}(t, \boldsymbol{p}) \in W$ for all $[0, \tau(\boldsymbol{p}))$ Corollary 6 can be applied to perform an identifiability analysis.

The smooth transformation λ, constructed via (9), is given by

$$\lambda(\boldsymbol{x}) = \left(\frac{p_1 + p_3 - \overline{p}_1 - \overline{p}_3 + \overline{p}_2 x_1}{p_2}, \frac{\overline{p}_7 x_2}{p_7}\right)^\top.$$

From the choice of μ_1 and (9) it is seen that λ automatically satisfies (12). It can be seen, by substituting for λ, that the last component of (11) is automatically satisfied. The first component can be rewritten in the form

$$c_1 + c_2 x_1(t, \overline{\boldsymbol{p}}) + c_3 x_2(t, \overline{\boldsymbol{p}}) + c_4 x_1(t, \overline{\boldsymbol{p}}) x_2(t, \overline{\boldsymbol{p}}) = 0 \quad (14)$$

where

$$c_1 = p_7(p_1(\overline{p}_1 + \overline{p}_3 - (p_1 + p_3)) + p_2 p_4) - \overline{p}_1 \overline{p}_2 \overline{p}_4),$$
$$c_2 = p_7(\overline{p}_1 - p_1)\overline{p}_2,\ c_3 = p_2(\overline{p}_1 + \overline{p}_3 - p_1 - p_3)\overline{p}_7,$$
$$c_4 = \overline{p}_2(p_7 \overline{p}_2 - p_2 \overline{p}_7).$$

Setting $t = 0$ in (14) and its first three derivatives with respect to t gives a system of equations that can be solved for the c_i to give $c_i = 0$ for all i. Solving this resulting set of four equations for the components of $\overline{\boldsymbol{p}}$ gives

$$\overline{p}_1 = p_1,\ \overline{p}_2 \overline{p}_4 = p_2 p_4,\ \overline{p}_3 = p_3,\ \overline{p}_7/\overline{p}_2 = p_7/p_2.$$

It now only remains to consider the initial condition (10). It is seen that

$$\lambda(\boldsymbol{x}_0(\overline{\boldsymbol{p}})) = \begin{pmatrix} p_5 \\ p_6 \end{pmatrix} = \begin{pmatrix} (\overline{p}_2 \overline{p}_5)/p_2 \\ (\overline{p}_7 \overline{p}_6)/p_7 \end{pmatrix}$$

and so $\overline{p}_2 \overline{p}_5 = p_2 p_5$ and $\overline{p}_6 \overline{p}_7 = p_6 p_7$. Therefore the set $\mathcal{S}(\boldsymbol{p})$ is given by

$$\left\{\overline{\boldsymbol{p}} \in \Omega : \overline{p}_1 = p_1, \overline{p}_3 = p_3, \overline{p}_2 \overline{p}_4 = p_2 p_4, \overline{p}_2 \overline{p}_5 = p_2 p_5, \overline{p}_2 \overline{p}_6 = p_2 p_6, \overline{p}_7/\overline{p}_2 = p_7/p_2\right\}$$

and, since the analysis holds for generic $\boldsymbol{p}$, the model is unidentifiable.

In terms of the original model parameters, the following are uniquely determined by the output:

- The parameters μ and σ;
- The combinations of parameters βN, $\beta S(0)$, $\beta I(0)$ and k/β.

This means that an uncountable number of different parameter vectors can give rise to the same output data. These parameter vectors correspond to particular choices for β. Note that the basic reproduction rate R_0 is globally identifiable since it is the ratio of βN and $\mu + \sigma$ which are both globally identifiable combinations.

If the size of the population, N, is known, then $\overline{p}_4 = p_4$ so that the set $\mathcal{S}(\boldsymbol{p})$ consists of the parameter vector $\boldsymbol{p}$ only. Similarly if any one of $S(0)$, $I(0)$, or k is known then $\mathcal{S}(\boldsymbol{p}) = \{\boldsymbol{p}\}$. In these cases the model is structurally globally identifiable.

When the (constant) effective contact rate β is replaced by the time-varying function defined in (5) then it is necessary to modify the model slightly in order to allow for the time-dependence of $\boldsymbol{f}$. Let $x_3 = \beta_1 \cos(2\pi t)$ and note that x_3 satisfies the following second order differential equation:

$$\ddot{x}_3 + 4\pi^2 x_3 = 0,$$

with initial conditions $x_3(0) = \beta_1$, and $\dot{x}_3(0) = 0$. The time-dependence of β, therefore, can be incorporated into the model by including two extra states, x_3 and x_4, that satisfy the following system of equations:

$$\begin{aligned} \dot{x}_3 &= x_4, & x_3(0) &= \beta_1, \\ \dot{x}_4 &= -4\pi^2 x_3, & x_4(0) &= 0. \end{aligned}$$

Note that, since $\cos(2\pi t) = x_3/\beta_1$, and $2\pi \sin(2\pi t) = (-x_4)/\beta_1$ are, in principle, known functions of time, the following outputs can also be included in the model:

$$y_1 = x_3/\beta_1 \qquad \text{and} \qquad y_2 = x_4/\beta_1.$$

Example 8. The general SIR model (1)–(2), together with the output (4) and seasonal forcing, modelled by the periodic (in time) function (5), can be written in the form

$$\begin{aligned} \dot{x}_1 &= p_1p_4 - x_1\left(p_1 + p_2x_2(1+x_3)\right) & x_1(0) &= p_5 \\ \dot{x}_2 &= x_2\left(p_2x_1(1+x_3) - (p_1+p_3)\right) & x_2(0) &= p_6 \\ \dot{x}_3 &= x_4 & x_3(0) &= p_7 \\ \dot{x}_4 &= -4\pi^2 x_3 & x_4(0) &= 0 \\ y_1 &= x_3/p_7 \\ y_2 &= x_4/p_7 \\ y_3 &= p_8x_2 \end{aligned}$$

where $x_1 = S$, $x_2 = I$, $p_1 = \mu$, $p_2 = \beta_0$, $p_3 = \sigma$, $p_4 = N$, $p_7 = \beta_1$, and $p_8 = k$. Let the vector of unknown parameters be denoted by $\boldsymbol{p} = (p_1, \ldots, p_8)^\top$ and the set of possible $\boldsymbol{p}$ be given by $\Omega = \{\boldsymbol{p} \in \mathbb{R}^8 : p_i > 0\ (i \neq 7),\ |p_7| < 1\}$. It is seen that $M(\boldsymbol{p}) = \mathbb{R}^4$.

To see that the ORC is satisfied at $\boldsymbol{x}_0(\boldsymbol{p})$ let $\mu_1(\boldsymbol{x}, \boldsymbol{p}) = x_3/p_7$, $\mu_2(\boldsymbol{x}, \boldsymbol{p}) = x_4/p_7$, $\mu_3(\boldsymbol{x}, \boldsymbol{p}) = p_8x_2$, and

$$\begin{aligned} \mu_4(\boldsymbol{x}, \boldsymbol{p}) &= \mathrm{L}_{\boldsymbol{f^p}}\mu_3^{\boldsymbol{p}}(\boldsymbol{x}) \\ &= p_8x_2\left(p_2x_1(1+x_3) - (p_1+p_3)\right). \end{aligned}$$

The Jacobian matrix of the function defined by $H_{\boldsymbol{p}}(\boldsymbol{x}) = (\mu_1(\boldsymbol{x}, \boldsymbol{p}), \ldots, \mu_4(\boldsymbol{x}, \boldsymbol{p}))^\top$ is given by

$$\frac{\partial H_{\boldsymbol{p}}}{\partial \boldsymbol{x}}(\boldsymbol{x}) = \begin{pmatrix} 0 & 0 & 1/p_7 & 0 \\ 0 & 0 & 0 & 1/p_7 \\ 0 & p_8 & 0 & 0 \\ p_2p_8x_2(1+x_3) & \alpha_1 & p_2p_8x_1x_2 & 0 \end{pmatrix}$$

where $\alpha_1 = p_8(p_2x_1(1+x_3) - (p_1+p_3))$. Given any $\boldsymbol{p} \in \Omega$ this matrix has full rank for all $\boldsymbol{x} \in W = \{\boldsymbol{x} \in \mathbb{R}^4 : x_2 \neq 0,\ x_3 \neq -1\}$. Since $\boldsymbol{x}(t, \boldsymbol{p}) \in W$ for all $[0, \tau(\boldsymbol{p}))$ Corollary 6 can be applied to perform an identifiability analysis.

The smooth transformation λ, constructed via (9), is given by

$$\lambda(\boldsymbol{x}) = \left(\lambda_1(\boldsymbol{x}), \frac{\overline{p}_8x_2}{p_8}, \frac{p_7x_3}{\overline{p}_7}, \frac{p_7x_4}{\overline{p}_7}\right)^\top$$

where

$$\lambda_1(\boldsymbol{x}) = \frac{\overline{p}_7\left(p_1 + p_3 - \overline{p}_1 - \overline{p}_3 + \overline{p}_2x_1(1+x_3)\right)}{p_2(\overline{p}_7 + p_7x_3)}.$$

Again, because of the choice of μ_1, equation (12) is automatically satisfied by λ. It can be seen, by substituting for λ, that the last three components of (11) are automatically satisfied. The first component can be rewritten in the form

$$\begin{aligned} &p_2p_8c_1x_2(t, \overline{\boldsymbol{p}})\alpha_2(t)^2 - p_8\big(c_6x_3(t, \overline{\boldsymbol{p}}) \\ &\qquad + c_3x_3(t, \overline{\boldsymbol{p}})^2 + \overline{p}_7(c_7 + p_7c_1x_4(t, \overline{\boldsymbol{p}}))\big) \\ &+ \overline{p}_2x_1(t, \overline{\boldsymbol{p}})\big(x_2(t, \overline{\boldsymbol{p}})\alpha_2(t)\alpha_3(t)(c_5 + c_5x_3(t, \overline{\boldsymbol{p}})) \\ &- p_8\overline{p}_7(c_2\overline{p}_7\alpha_3(t) + c_2p_7x_3(t, \overline{\boldsymbol{p}}) + p_7c_2x_3(t, \overline{\boldsymbol{p}})^2 \\ &\qquad\qquad + c_4x_4(t, \overline{\boldsymbol{p}}))\big) = 0 \end{aligned} \tag{15}$$

where

$$\begin{gathered} \alpha_2(t) = \overline{p}_7 + p_7x_3(t, \overline{\boldsymbol{p}}), \quad \alpha_3(t) = 1 + x_3(t, \overline{\boldsymbol{p}}), \\ c_1 = (\overline{p}_1 + \overline{p}_3) - (p_1 + p_3), \quad c_2 = p_1 - \overline{p}_1, \\ c_3 = p_7(\overline{p}_1\overline{p}_2\overline{p}_4\overline{p}_7 - p_1p_2p_4p_7), \\ c_4 = \overline{p}_7 - p_7, \quad c_5 = \overline{p}_2\overline{p}_7p_8 - p_2\overline{p}_7\overline{p}_8, \\ c_6 = \overline{p}_7\big(p_1p_7(c_1 - 2p_2p_4) + \overline{p}_1\overline{p}_2\overline{p}_4(p_7 + \overline{p}_7)\big), \\ c_7 = \overline{p}_7\big(p_1(c_1 - p_2p_4) + \overline{p}_1\overline{p}_2\overline{p}_4\big). \end{gathered}$$

Setting $t = 0$ in (15) and its first six derivatives, with respect to t, gives a system of equations that can be solved for the c_i to give $c_i = 0$ for all i. Solving this resulting set of seven equations for the components of $\overline{\boldsymbol{p}}$ gives

$$\begin{gathered} \overline{p}_1 = p_1, \quad \overline{p}_2\overline{p}_4 = p_2p_4, \quad \overline{p}_3 = p_3, \\ \overline{p}_8/\overline{p}_2 = p_8/p_2, \quad \overline{p}_7 = p_7. \end{gathered}$$

It now only remains to consider the initial condition (10). It is seen, since $\lambda_1(\boldsymbol{x}) = (\overline{p}_2x_1)/p_2$, that

$$\begin{aligned} \lambda(\boldsymbol{x}_0(\overline{\boldsymbol{p}})) &= \left(p_5, p_6, p_7, 0\right)^\top \\ &= \left((\overline{p}_2\overline{p}_5)/p_2, (\overline{p}_2\overline{p}_6)/p_2, p_7, 0\right)^\top \end{aligned}$$

and so $\overline{p}_2\overline{p}_5 = p_2p_5$ and $\overline{p}_2\overline{p}_6 = p_2p_6$. Therefore the set $\mathcal{S}(\boldsymbol{p})$ is given by

$$\begin{gathered} \{\overline{\boldsymbol{p}} \in \Omega : \overline{p}_1 = p_1, \overline{p}_3 = p_3, \overline{p}_7 = p_7, \overline{p}_2\overline{p}_4 = p_2p_4, \\ \overline{p}_2\overline{p}_5 = p_2p_5, \overline{p}_2\overline{p}_6 = p_2p_6, \overline{p}_8/\overline{p}_2 = p_8/p_2\} \end{gathered}$$

and, since the analysis holds for generic $\boldsymbol{p}$, the model is unidentifiable.

In terms of the original model parameters, the following are uniquely determined by the output:

- The parameters μ, σ and β_1;
- The combinations of parameters β_0N, $\beta_0S(0)$, $\beta_0I(0)$ and k/β_0.

It is interesting to note that the amplitude of the seasonal variation, β_1, is globally identifiable, but the average value, β_0, is unidentifiable. Therefore, for a given output, $\beta_0 > 0$ is free for choice, with the other parameters being chosen to satisfy the relations in $\mathcal{S}(\boldsymbol{p})$.

As in Example 7, if the size of the population, N, is known, $\overline{p}_4 = p_4$ and so $\mathcal{S}(\boldsymbol{p}) = \{\boldsymbol{p}\}$. This is also true if any one of $S(0)$, $I(0)$, or k is known. In these cases the model is structurally globally identifiable.

One might also wish to test for the structural identifiability of the model (1)–(3) in the case when a proportion of the rate of incidence is measured. This corresponds to a measurement of the number of newly infected individuals and can be modelled by the inclusion of an output of the form

$$y = kSI$$

in the model. This output could also be allowed to vary seasonally as the force of infection, which gives rise to an output of the form

$$y = k(1 + \beta_1 \cos(2\pi t))SI.$$

At present the extra complexity in the identifiability analysis caused by the inclusion of outputs of these forms, is beyond the computational bounds of the symbolic package *Mathematica* (Wolfram, 1999) that was used for the above examples. This is a topic of ongoing research.

5. CONCLUSIONS

It has been shown that a general SIR model, with the force of infection subject to seasonal variation, is unidentifiable when a proportion of the number of infectives is measured. This type of output is appropriate for models of diseases with a short period of infection, such as respiratory syncytial virus (Weber *et al.*, 2001). However, if the size of the total population is known, or it is known what proportion of the number of infectives is measured, then the model is structurally globally identifiable.

The methodology employed in this paper for the identifiability analysis is essentially that developed by Evans *et al.* (2002). Small modifications to the model were necessary to apply the theory when the effective contact rate was subject to temporal variation. This approach utilises the existence of a smooth transformation connecting the state trajectories corresponding to indistinguishable parameter vectors. As such it is similar to the similarity transformation approach for controlled parametric models (Vajda *et al.*, 1989).

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the Engineering and Physical Sciences Research Council of the UK under grant GR/M11943.

6. REFERENCES

Bellman, R. and K. J. Åström (1970). On structural identifiability. *Math. Biosci.* **7**, 329–339.

Capasso, V. (1993). *Mathematical Structures of Epidemic Systems*. Springer-Verlag. Berlin.

Chappell, M. J., K. R. Godfrey and S. Vajda (1990). Global identifiability of the parameters of nonlinear systems with specified inputs: A comparison of methods. *Math. Biosci.* **102**, 41–73.

Evans, N. D., M. J. Chapman, M. J. Chappell and K. R. Godfrey (2002). Identifiability of uncontrolled nonlinear rational systems. *Automatica*. Provisionally accepted.

Fliess, M. and S. T. Glad (1993). An algebraic approach to linear and nonlinear control. In: *Essay on Control: Perspectives in the Theory and its Applications* (H. L. Trentelman and J. C. Willems, Eds.). Vol. 14 of *Progress in Systems and Control Theory*. Birkhäuser, Boston.

Godfrey, K. R. and J. J. DiStefano III (1987). Identifiability of model parameters. In: *Identifiability of Parametric Models* (E. Walter, Ed.). Chap. 1, pp. 1–20. Pergamon Press. Oxford.

Grenfell, B., B. Bolker and A. Kleczkowski (1995). Seasonality, demography and the dynamics of measles in developed countries. In: *Epidemic models: their structure and relation to data* (D. Mollison, Ed.). pp. 248–268. Publications of the Newton Institute. Cambridge University Press. Cambridge.

Hermann, R. and A. J. Krener (1977). Nonlinear controllability and observability. *IEEE Trans Automat. Control* **22**, 728–740.

Kermack, W. O. and A. G. McKendrick (1927). Contributions to the mathematical theory of epidemics, Part I. *Proc. Roy. Soc., A* **115**, 700–721.

Ljung, L. and T. Glad (1994). On global identifiability for arbitrary model parametrizations. *Automatica* **30**, 265–276.

Pohjanpalo, H. (1978). System identifiability based on the power series expansion of the solution. *Math. Biosci.* **41**, 21–33.

Tunali, E. T. and T. J. Tarn (1987). New results for identifiability of nonlinear systems. *IEEE Trans. Automat. Control* **32**, 146–154.

Vajda, S., K. R. Godfrey and H. Rabitz (1989). Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Math. Biosci.* **93**, 217–248.

Walter, E., (Ed.) (1987). *Identifiability of Parametric Models*. Pergamon Press. Oxford.

Weber, A., M. Weber and P. Milligan (2001). Modeling epidemics caused by respiratory syncytial virus (RSV). *Math. Biosci.* **172**, 95–113.

Wolfram, S. (1999). *The Mathematica Book*. 4th ed. Wolfram Media/Cambridge University Press. Cambridge, UK. Mathematica Version 4.