Environmental Modelling & Software 93 (2017) 398–408

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Assessing local structural identifiability for environmental models

J.D. Stigter [a, *], M.B. Beck [b], J. Molenaar [a]

[a] Biometris, Wageningen University and Research, Wageningen, The Netherlands
[b] Department of Civil and Environmental Engineering, Imperial College, South Kensington Campus, London, United Kingdom

ARTICLE INFO

Article history:
Received 27 November 2015
Received in revised form
7 March 2017
Accepted 7 March 2017

Keywords:
Identifiability
Non-linear parameter estimation
Parametric output sensitivity
Model reduction

ABSTRACT

The local structural identifiability problem is investigated for the general case and demonstrated for a well-known microbial degradation model that includes 13 unknown parameters and 3 additional states. We address the identifiability question using a novel algorithm that can be used for large models with many parameters to be identified. A key ingredient in the analysis is the application of a singular value decomposition of the normalized parametric output sensitivity matrix that is obtained through a simple model integration. The SVD results are further analysed and verified in a complementary symbolic computation. It is especially the swiftness and accuracy of the suggested method that we consider to be a substantial advantage in comparison to existing methods for a structural identifiability analysis. The method also opens, in a natural way, the analysis of (parametric) uncertainty in general, and this is demonstrated in more detail in the results section.

1. Introduction

In environmental modelling local structural identifiability is an essential model property that should be investigated as part of a calibration of the model against field data. Structural identifiability may even be investigated *before* the data have been collected. Beck already writes in his 1987 review on the analysis of uncertainty in water quality models that 'For water quality models the overriding difficulty of parameter estimation is that of a lack of parameter identifiability', Beck (1987). As for the definition of this fundamental problem a distinction can be made between structural or theoretical identifiability (assuming perfect data) and practical identifiability (assuming noise corrupted data). In the former problem, structural identifiability amounts to the question: Is it possible to have *exact equivalence of two output signals*, say $y(t,\theta_1)$ and $y(t,\theta_2)$, that have been generated using two *different* sets of parameter values θ_1 and θ_2? Or, put differently, does equivalence of the two output signals $y(t,\theta_1) = y(t,\theta_2)$ imply equivalence of the parameter values that generated these two curves, i.e. $\theta_1 = \theta_2$? For practical identifiability the main question in general is how many parameters can be uniquely identified on the basis of *real* measurement/output signals that are inevitably corrupted with noise. Clearly, theoretical identifiability is a prerequisite for practical identifiability. We refer to Cobelli and DiStefano III (1980); Norton (1980); Walter (1982); Godfrey (1986) for a more detailed discussion on structural identifiability. We also note that structural identifiability can be assessed both locally and globally. If a model is not globally identifiable, distinct sets of parameter values yield exactly the same output. For example, if the parameter value $\theta_1 = 1$ must be determined from sensor readings, while in the model equations θ_1^2 is the only expression that includes this parameter, then clearly $\theta_1 = -1$ yields the same output signal as $\theta_1 = 1$. Therefore θ_1 is not globally identifiable in this case, while locally it is.

In a more recent contribution various aspects of identifiability for environmental models are reviewed in Marsili-Libelli et al. (2014). A more technical paper that summarizes the various methods to assess structural identifiability is Miao et al. (2011), in which a good and complete overview is presented. Here, we consider *local* structural identifiability for the general and well-known non-linear state-space model that reads

$$\frac{dx(t)}{dt} = f(x(t), u(t), \theta) \tag{1}$$

$$y(t) = h(x(t), u(t), \theta) \tag{2}$$

where $x(t)$ is a vector of states ($\dim(x) = n$), $u(t)$ is a vector of inputs

* Corresponding author. Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
E-mail address: hans.stigter@wur.nl (J.D. Stigter).

http://dx.doi.org/10.1016/j.envsoft.2017.03.006
1364-8152/

($\dim(u) = r$), θ is the vector of parameters ($\dim(\theta) = p$), and $y(t)$ is the vector of model outputs (or the observation signals, $\dim(y) = m$). Although the method presented in this paper can be extended to more general model types, we focus our discussion on non-linear state space models. An important assumption that is necessary for the symbolic part of our identifiability algorithm to work, is that the vector valued functions f and h are *smooth*, meaning that their derivatives with respect to the states $x(t)$ and parameters θ exist. This excludes, in principle, the situation where a switch in the model definition is included that limits the model to a certain domain such as, for example, in the Blackman growth model that describes the growth rate of a microbial population (and which may be part of the non-linear model definition):

$$r(x, \theta_1) = \begin{cases} \theta_1 x, & \theta_1 x \le \theta_2 \\ \theta_2, & \theta_1 x > \theta_2 \end{cases}$$

This type of discontinuity, however, can easily be remedied using a *smooth switching function* that allows the derivatives of the output $y(t) = h(x(t), u(t), \theta)$ to be calculated symbolically at any time instant t with the help of an algebraic software package. In the results section of this paper we will present another example of a discontinuity in the model structure f due to a time-lag in the modelled process that causes non-smoothness. This will, again, be solved with the aid of a smooth switching function that guarantees the derivatives to exist at any time instant t. Alternatively, one may study identifiability for both cases of the 'if-then' statement in the model structure. For the Blackman model this means we study identifiability in two separate regions in the state-space, one for $\theta_1 x \le \theta_2$, and one for $\theta_1 x > \theta_2$. The outcome of the identifiability analysis may then be different for the two cases but, of course, the two separate analyses yield complementary results that are certainly of interest before the actual parameter estimation is performed.

Brun et al. (2001), in their paper on practical identifiability of large environmental models, present a framework that allows one to answer, in principle, the practical identifiability question for the general model (1)–(2). It is not their intention to assess theoretical identifiability in their examples. Rather, the focus in their work is on finding identifiable sets of parameters that can be reliably estimated given an experimental layout and corresponding data set. This question is resolved on the basis of an analysis of the parametric output sensitivities. Here, we study *both* theoretical and practical identifiability for the general non-linear model structure (1)–(2), with most emphasis on the theoretical identifiability question. Our approach is based on an analysis of the (dynamic) parametric output sensitivities that describe the change of the outputs $y(t)$ caused by perturbations in the parameters θ. The sensitivity dynamics can be easily derived from model (1)–(2) as:

$$\frac{d}{dt}\frac{\partial x(t,\theta)}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial \theta} \tag{3}$$

$$\frac{\partial y(t,\theta)}{\partial \theta} = \frac{\partial h}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial h}{\partial \theta} \tag{4}$$

Note that $\frac{\partial x(t,\theta)}{\partial \theta}$ is a $(n \times p)$ matrix, whose columns contain the sensitivity of each state-vector element $x_i(t), i = 1, \ldots, n$, to one specific parameter θ_j, $j = 1, \ldots, p$ in the parameter vector θ. To guarantee accuracy of our results, all Jacobi matrices $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial \theta}$, $\frac{\partial h}{\partial x}$, and $\frac{\partial h}{\partial \theta}$ are calculated symbolically. Once the combined equations (1)–(4) have been solved numerically for a specific (also referred to as nominal) vector $\bar{\theta}$, and the solution is known on a time grid $[t_0, \ldots, t_N]$, we build the (dimensionless) normalized sensitivity matrix $S_n(t_0, \ldots, t_N, \bar{\theta})$:

$$S_n\left(t_0, \ldots, t_N, \bar{\theta}\right) = \begin{pmatrix} \frac{\bar{\theta}_1}{y_1(t_0)}\frac{\partial y_1(t_0)}{\partial \theta_1} & \cdots & \frac{\bar{\theta}_p}{y_1(t_0)}\frac{\partial y_1(t_0)}{\partial \theta_p} \\ \vdots & & \vdots \\ \frac{\bar{\theta}_1}{y_m(t_0)}\frac{\partial y_m(t_0)}{\partial \theta_1} & \cdots & \frac{\bar{\theta}_p}{y_m(t_0)}\frac{\partial y_m(t_0)}{\partial \theta_p} \\ \vdots & & \vdots \\ \frac{\bar{\theta}_1}{y_1(t_N)}\frac{\partial y_1(t_N)}{\partial \theta_1} & \cdots & \frac{\bar{\theta}_p}{y_1(t_N)}\frac{\partial y_1(t_N)}{\partial \theta_p} \\ \vdots & & \vdots \\ \frac{\bar{\theta}_1}{y_m(t_N)}\frac{\partial y_m(t_N)}{\partial \theta_1} & \cdots & \frac{\bar{\theta}_p}{y_m(t_N)}\frac{\partial y_m(t_N)}{\partial \theta_p} \end{pmatrix} \tag{5}$$

where we have assumed that the output signals $\{y_j(t), j = 1, \ldots, m\}$ are not equal to zero for any time t.[1] The relative sensitivity matrix can easily be evaluated numerically for many random vectors $\{\bar{\theta}^i, i = 1, 2, \ldots\}$. To test for local structural identifiability of the model, we evaluate the rank of $S_n(t_0, \ldots, t_N, \bar{\theta})$ using a singular value decomposition. We assume that the random vector $\bar{\theta}$ at which the sensitivities are evaluated is *regular*[2], meaning that the rank of $S_n(t_0, \ldots, t_N, \bar{\theta})$ does not change within a small neighbourhood of the nominal parameter vector $\bar{\theta}$. The reason for imposing regularity of the parameter vector $\bar{\theta}$ is that, for example, in the simple model

$$\frac{dx(t)}{dt} = -\theta_1 x(t) \tag{6}$$

$$x(t_0) = \theta_0 \tag{7}$$

$$y(t) = x(t) \tag{8}$$

the parameter θ_1 clearly can *not* be identified if we set $\theta_0 = 0$, while θ_0 can be estimated by simple observing the state at $t = 0$. Clearly, a small perturbation on $\theta_0 = 0$ changes the outcome of a structural identifiability analysis, i.e. $\theta_0 = 0$ is a non-regular point, and we prefer to focus the discussion on regular points since these are the most likely ones to be encountered.

Furthermore, when building the matrix $S_n(t_0, \ldots, t_N, \bar{\theta})$, the time grid $[t_0, \ldots, t_N]$ on which the local sensitivities are solved is carefully chosen, since it will influence the accuracy of our rank test. Accuracy of the SVD algorithm in general is discussed in the book by Golub and van Loan (1996), section 5.5.8. For the matrix $S_n(t_0, \ldots, t_N, \bar{\theta})$, this analysis implies that for an accurate outcome of the SVD algorithm, the time grid is chosen in such a way that the number of rows in the matrix $S_n(t_0, \ldots, t_N, \bar{\theta})$ is as small as possible, but not smaller than the number of parameters p. Hence, we choose $m\,(N+1) > p$ with N as small as possible.

The key message we want to convey in this paper is that to

[1] To remedy possible "zeroness" of the outputs, we can scale the output with an average $\bar{y}_j(t)$ to obtain non-zero denominators in the matrix entries in equation (5). Even in case the output $y(t)$ is very small, a scaling factor can be applied to remedy large values of the relative sensitivities.

[2] We refer to Kwatny and Blankenship (2000) for a more elaborate treatment of regularity of the point $\bar{\theta}$. In that reference this property is important when establishing controllability and/or observability properties of a non-linear control-system on the basis of a rank test.