University of Chester

Author(s): Christopher T H Baker ; Gennady Bocharov ; F A R Rihan

Title: Neutral delay differential equations in the modelling of cell growth

Date: 2008

Example citation: Baker, C. T. H., Bocharov, G., & Rihan, F. A. R. (2008). *Neutral delay differential equations in the modelling of cell growth* (Applied Mathematics Group Research Report, 2008 : 1). Chester, United Kingdom: University of Chester.

Version of item: Published version

Available at: http://hdl.handle.net/10034/346421

University of Chester

Department of Mathematics

Applied Mathematics Group
Research Report

2008 Series

NEUTRAL DELAY DIFFERENTIAL EQUATIONS IN THE MODELLING OF CELL GROWTH

Christopher T H Baker, Gennady A Bocharov, & Fathalla A Rihan

NEUTRAL DELAY DIFFERENTIAL EQUATIONS IN THE MODELLING OF CELL GROWTH

Christopher T H Baker[a,b] *,

[a] *Department of Mathematics, University of Chester, Chester CH1 4BJ;*
[b] *Emeritus professor, The University of Manchester, M13 9PL, UK.*

Gennady A. Bocharov [c 1],

[c] *Institute of Numerical Mathematics, Russian Academy of Science, Moscow*

Fathalla A. Rihan [d 2]

[d] *Department of Mathematics, Faculty of Science, Helwan University, Egypt*

Abstract

In this contribution, we indicate (and illustrate by example) rôles that may be played by *neutral delay differential equations* in modelling of certain cell growth phenomena that display a time lag in reacting to events. We explore, in this connection, questions involving the sensitivity analysis of models and related mathematical theory; we provide some associated numerical results.

Key words: Computational modelling, Delay & neutral delay-differential equations, Models of growth phenomena, Parameter estimation, Variation of parameters, Sensitivity, Model selection.
2000 MSC (AMS Mathematics Subject Classification): 93A13, 93A30, 34K05, 34K28, 34K40, 45G15, 47A05, 47N20, 47N60, 62P10, 90C31

1 Introduction

The problem of formulating a mathematical model explaining the systems behaviour specified by time-series of observations is considered to be central to every scientific discipline. Authors of a number of papers or books have addressed the determination of computational models that describe mathematically the evolution of biological

---

* Corresponding author, *E-mail contact:* `cthbaker@na-net.ornl.gov`.

[2] Currently located at: Department of Mathematical Sciences, United Arab Emirates University, Al Ain, Abu Dhabi

phenomena. In general, such phenomena depend upon past states (phenomena that involve a gestation period are possibly the first examples to come to mind). Below, we indicate, and illustrate by considering an example of cell growth dynamics, rôles that may be played by *neutral delay differential equations* (NDDEs) in modelling of certain growth phenomena that display a time lag in their reaction to events.

By way of orientation, the NDDEs that are discussed here are of the form

$$\frac{d}{dt}\Big\{y(t) - g_0\big(\mathbf{p}; t, y(t), y(t - \tau)\big)\Big\} = g_1\big(\mathbf{p}; t, y(t), y(t(t - \tau))\big) \quad (t \geq t_0), \quad (1.1)$$

or alternatively

$$y'(t) = g\big(\mathbf{p}; t, y(t), y(t - \tau), y'(t - \tau)\big) \quad (t \geq t_0), \quad (1.2)$$

in each case incorporating a parameter vector $\mathbf{p} = [p_1, p_2, \cdots, p_L]^T \in \mathbb{R}^L$ and with $\tau > 0$. τ may be one of the components of $\mathbf{p}$; exceptionally, we might admit the possibility $\tau = 0$. To determine a solution $y(\cdot) = y(\mathbf{p}; \cdot)$ of such [iii] an NDDE we require an initial condition of the form $y(t) = \psi(t)$ for $t \in [t_0 - \tau, t_0]$.

1.1 Our perspective

For the present authors, a mathematical model comprises a set of equations, and possibly constraints, that have their foundations in scientific observations and theory and purport to generate insight into some specified phenomena. Our perspective is that we should provide quantitative as well as qualitative predictions, and the ultimate models are therefore computational models.

Rather than advocating a single, definitive model of biological growth phenomena, we consider a variety of modelling approaches that have different strengths, weaknesses, and domains of applicability. The models with which we concern ourselves are based on ordinary differential equations [ODEs], delay-differential equations [DDEs], and NDDEs. These are all examples of retarded functional differential equations (RFDEs). (In practice, complex models containing equations of various types are to be combined.) It is from the modelling context that we approach our present topic. Note that improvements in modelling precision typically carry a cost: they may require more specific or precise data to be properly calibrated. We concentrate on features involving NDDEs but, for intelligibility, we develop some of the background.

[iii] In Remark 1.1, we indicate how to normalize (if required) so that $t_0 = 0$ and $\tau > 0$ is replaced by 1.

1.2 Example equations

To open the discussion, note that a class of linear NDDE comprises equations of the form

$$\{y(t) - \rho_2 y(t - \tau)\}' = \rho_0 y(t) + \rho_1 y(t - \tau) \quad (t \in [t_0, T)) \text{ where } \tau > 0. \quad (1.3a)$$

If the solution $y(\cdot)$ is differentiable, this equation can be rewritten as

$$y'(t) = \rho_0 y(t) + \rho_1 y(t - \tau) + \rho_2 y'(t - \tau). \quad (1.3b)$$

It is convenient to associate the equation with a parameter $\mathbf{p} = [\rho_0, \rho_1, \rho_2]^T$ or $\mathbf{p} = [\rho_0, \rho_1, \rho_2, \tau]^T$ (but see Remark 1.1, below).

For either equation (1.3a) or (1.3b), we require an initial function ψ (which may itself depend on parameters, say $\psi(t) \equiv (\psi(\mathbf{p}; t))$ to specify a solution, and we write

$$y(t) \equiv y(\psi, \mathbf{p}; t) \quad (1.3c)$$

where

$$y(\psi, \mathbf{p}; t) = \psi(t) \text{ for } t \in [t_0 - \tau, t_0]. \quad (1.3d)$$

We designate by $\mathbf{y}_0$ the value

$$\mathbf{y}_0 = \psi(t_0). \quad (1.3e)$$

From time to time we suppress arguments so that

$$y(t) \text{ is identified with } y(\mathbf{p}, t), \text{ or } y(\psi, \mathbf{p}; t), \text{ or } y(\psi; t), \text{ according to context.} \quad (1.3f)$$

Unless otherwise stated it is assumed that $\psi : [t_0 - \tau, t] \to \mathbb{R}$ has a derivative (either right-hand [iv] or conventional) and [v]

$$y'(\psi, \mathbf{p}; t) = \psi'(t) \text{ for } t \in [t_0 - \tau, t_0]. \quad (1.3g)$$

The smoothness of the solution $y(\psi; t)$ is determined by ψ and it should be noted that a solution of an NDDE can be discontinuous [10,12] but more generally it is continuous with a discontinuous derivative. For y to be differentiable at t_0 we require the left-hand derivative $\psi'_-(t_0)$ to match the right-hand derivative $y'_+(t_0)$ determined by the NDDE at $t = t_0$. Henceforth, t_0 is prescribed and we regard $\mathbf{y}_0$ as either known from data or unknown (see §2.2).

[iv] The right-hand derivative is written, where we need to distinguish it, as ψ'_+. The value $\psi'(t_0)$ is the left-hand derivative $\psi'_-(t_0)$ and the value $\psi'(t_0 - \tau)$ is the right-hand derivative, $\psi'_+(t_0 - \tau)$.

[v] In certain modelling conditions $y'(\psi; t) = \psi'(t)$ is not assumed to be the case and we replace (1.3g) by an independent condition $y'(t) = \psi_1(t)$ $(t \in [t_0 - \tau, t_0])$.

A class of DDEs arises if we restrict attention to the case $\rho_2 = 0$ and a class of ODEs arises if we further require $\rho_1 = 0$. Thus, the corresponding ODEs, DDEs, and NDDEs form a nested hierarchy of models:

$$y'(t) = \rho_0 y(t) \qquad\qquad (t \geq t_0), \tag{1.4a}$$
$$y'(t) = \rho_0 y(t) + \rho_1 y(t - \tau) \qquad (t \geq t_0), \tag{1.4b}$$
$$y'(t) = \rho_0 y(t) + \rho_1 y(t - \tau) + \rho_2 y'(t - \tau) \qquad (t \geq t_0). \tag{1.4c}$$

A case has been presented [3,5] for a model based on the equation

$$y'(t) = \frac{1}{\tau} y(t - \tau) \quad (t \geq t_0). \tag{1.4d}$$

This equation is a special case of our standard equation (1.3) (see also (1.4b)) but with

$$\rho_0 = 0, \ \rho_1 = \tau \ (\text{equivalently } \tau = \rho_1), \ \text{and } \rho_2 = 0.$$

Remark 1.1 *Corresponding inhomogeneous equations (often termed* affine *NDDEs), have the form e.g.,*

$$\{y(t) - \rho_2 y(t - \tau)\}' = \rho_0 y(t) + \rho_1 y(t - \tau) + f(t) \quad (t \in [t_0, T)).$$

Observe that by a change of variable ($t := \tau s + t_0$) one can normalize $\tau > 0$ to have the value unity and t_0 to have the value 0. For example, the previous equation is equivalent to

$$u'(s) = \{\tau \rho_0\} u(s) + \{\tau \rho_1\} u(s - 1) + \rho_2 u'(s - 1) + \tau f(\tau s + t_0) \quad for \ s \geq 0, \quad (1.5)$$

where $u(s) = y(\tau s + t_0)$ for $s \geq -1$. In principle, the introduction of τ as a parameter may be considered via such a transformation.

1.3 Conservation laws

It may be asked from whence the NDDEs of the above form originate. One source is the hyperbolic equations of conservation law models in structured population dynamics. The corresponding models are associated with the names of Lotka-Sharpe-McKendrck and Gurtin-MacCamy. Consider, as a simple example, the linear Sharpe-Lotka-McKendrick model for a population structured by age a and specified at time t by the distribution function $n(t, a)$ (see the analysis presented in detail in [22]):

$$n_t(t, \alpha) + n_\alpha(t, \alpha) + \mu(\alpha) n(t, \alpha) = 0$$

with $n(t, 0) = \int_0^\infty b(\alpha) n(t, \alpha) d\alpha$, $n(0, \alpha) = n_0(\alpha)$. The birth $b(\alpha)$ and the death $\mu(\alpha)$ rate functions depend on age. Assume that there is a critical age $\tau > 0$ which separates adult age classes $\alpha > \tau$ from juvenile age classes $\alpha < \tau$. Then, one can introduce the total juvenile $z(t)$ and the total adult population $y(t)$ as $z(t) = \int_0^{\tau^-} n(t, a) da,$

and $y(t) = \int_{\tau+}^{\infty} n(t, a) da$. Further assume that $b(\alpha) = (\rho_1 + \rho_2 \rho_0) H_\tau(\alpha) + \rho_2 \delta_\tau(\alpha)$, and $\mu(\alpha) = -\rho_0$, where $H_\tau(\alpha)$ is a Heaviside function with jump at $\alpha = \tau$ and $\delta_\tau(\alpha)$ is a delta function. Then one can show that the function $y(t)$ satisfies a linear NDDE

$$y'(t) = \rho_0 y(t) + \rho_1 y(t - \tau) + \rho_2 y'(t - \tau). \tag{1.6}$$

1.4 Existing theory & practice

The basic theory for NDDEs is quite well covered in the literature (for example, [33,34]) and we shall be selective in the remarks offered below, concentrating on features that are particularly relevant to us.

Let us first consider briefly the *interpretation of the neutral delay equation*. It is possible to consider solutions of NDDES (a) in the sense of Carathéodory, [10,27] (b) in the sense of one-sided derivatives, (c) in the sense of conventional derivatives. With each interpretation of the equation one can enquire whether a solution $y(t)$ is, or is *required* to be, (i) piecewise continuous, (ii) piecewise continuous and piecewise differentiable, (iii) continuous, (iv) continuous and differentiable in the conventional sense on $[t_0, T)$, and so on. We shall comment below on our cell-growth model, the initial conditions, and the sense in which we consider solutions to be defined. The smoothness of the solution is an important (and related) issue in modelling. To match the differentiability conditions, and given a sufficiently smooth function v, we can write

$$L_+\{v\}(t) := v'_+(t) - \{\rho_0 v(t) + \rho_1 v(t - \tau) + \rho_2 v'_+(t - \tau)\} \tag{1.7a}$$

when v has a right-hand derivative at t and $t - \tau$ and

$$L\{v\}(t) := v'(t) - \{\rho_0 v(t) + \rho_1 v(t - \tau) + \rho_2 v'(t - \tau)\} \tag{1.7b}$$

when v has a conventional derivative. Since (1.7a) collapses to (1.7b) whenever the two-sided derivative exists, we can still use the notation L_+ in (1.7a) in some places where the notation (1.7b) would be more transparent.

Smoothness of the solution is also relevant in the numerics [2,20] of DDEs and NDDEs (the solution or its derivatives may inherit jumps from the behaviour of the initial function, through the lag). In consequence, a distinction between two-sided and right-hand derivatives may be appropriate. Possible assumptions on $\psi(t)$ include that ψ be (a) bounded (b) piecewise-continuous, (c) piecewise-continuous and differentiable from the right, (d) continuous and differentiable, or (e) continuously differentiable.

Remark 1.2 *Additional background literature can be categorized as follows:*

(1) Classical analysis of RFDEs – see, e.g.[21,31,34,35]
(2) Modelling, notably in bioscience; see, e.g. [38] and [1,3,6,4,5,8,9,23,19]
(3) Numerics of RFDEs [2,12,20,47]
(4) Computational codes for various types of RFDE [30,39,42]

1.5   The structure of the paper

Following this introduction, we have a section on modelling considerations, followed by a section on perturbation theory, one on sensitivity, another on generalizations and future work, and we conclude with references.

2   Modelling considerations

In this paper we concentrate on NDDEs, and their application is illustrated by reference to experimental data reflecting the growth patterns of cells which clearly exhibit the presence of a time-lag in their division, e.g., the synchronous *in vitro* culture of *Escheria coli* (*E.coli*). Given such data and a specified type of formal model we seek an actual model in which the parameters (such as $\{\rho_\ell\}$ and τ above) that define the structure are assigned values that are matched to the data. This process can be set within the larger context of selecting a best parameterized model chosen from within a hierarchy according to principles based on parsimony, Bayesian concepts, maximum likelihood or minimimum description length; see [24,29,37] *etc.*

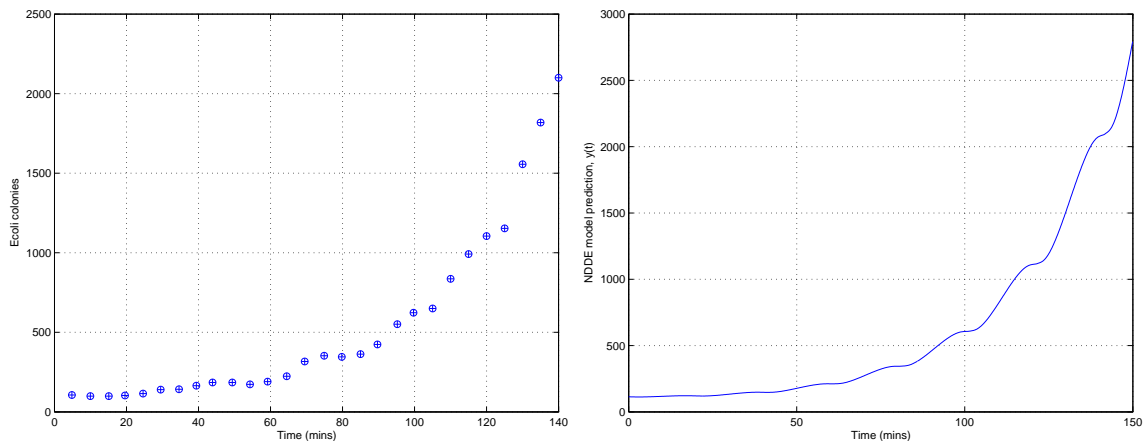––-

Fig. 1. These figures show (left) data for synchronous *E.coli* growth (based on [45, Fig.4]) and (right) a graph of a solution of an NDDE (1.3), with particular values of its parameters and initial function. The two figures show that an NDDE model with appropriate parameters can be qualitatively consistent with the observed growth of synchronous cell populations.

––-

6

2.1 A first look at a modelling example

Given experimental data on cell growth, the use of such models as those above may enable modellers to estimate some relevant growth parameters such as: the cell-doubling time, the fraction of cells that are dividing, the rate of commitment of cells to cell division, the degree of synchronization of cells in the population, and the death rate of cells. Figure 1 shows observed data for synchronous *E.coli* growth (based on [45, Fig.4]) compared with the graph of a solution $y(t)$ of the NDDE (1.3b) that corresponds to particular values of the parameters and an initial function. The curve is clearly consistent with the observed data and illustrates the stepwise growth of the population.

We shall use this modelling example as the basis for illustrating the general results available in the context of model construction, analysis, and selection.

2.2 Cell growth and observational data

Cell growth is a process central to every living system. A cell population is an ensemble of individual cells, all of which contribute in a different way to the overall observed behaviour. The heterogeneity of dividing cell populations can be described by a wide range of phenotypic and/or physical characteristics, *e.g.*, the doubling time, the position in the cell cycle. Various proliferation assays are used to quantify the turnover of growing cell populations. Cell growth kinetics is affected by the cellular heterogeneity with respect to the division rate, initial position in the cell cycle, etc. Broadly heterogeneous cell populations display an exponential net growth pattern (for as long as the necessary resources are available), whereas initially synchronized cell cultures with a low variability in their division characteristics show for some time a non-monotone step-like growth until they reach a 'steady exponential growth'. A representative example of such a non-monotone growth is provided by the *in vitro* system of synchronized *E.coli* cells [45]. It represents a nicely understood reference growth system in bioengineering research (see Chapter 6.5.1 in [15]). We used the published data on the population size of *E.coli* as a function of time to estimate the parameters of the above nested family of models. In addition, as there was no statistical characterization of the measurement error in the original study, we estimated the standard deviation in the maximum likelihood framework.

2.3 Cell growth and parameter selection

The NDDE (1.3) contains parameters $\rho_{0,1,2}, \tau$; the initial conditions (see below) involve further parameters. In general, the issue is to choose a model in which the parameters and the initial conditions are scientifically meaningful and determine "suitable" numerical values of these parameters.

We take $t_0 = 0$ here and in the subsequent discussion of cell growth so that the

model equations (1.3) read

$$\mathsf{L}\{y\}(t) = 0 \ (t \geq 0), \text{ and } \mathsf{L} \text{ is one of (1.7a) or (1.7b)}. \tag{2.1a}$$

We settle the choice of L from (1.7a) or (1.7b) (equivalently, the interpretation of the model equations (2.1a)) in §2.4. Clearly, in either case L depends on $\{\rho_\ell\}, \tau$.

Introducing further parameters ρ_3 and $\rho_\star$ ($\rho_\star$ will later be absorbed in ρ_3), we define the initial function $\psi(t)$ for $t \in [-\tau, 0]$ by defining $\psi(0) = \mathsf{y}_0$ (which is either a measured observation or has to be estimated as a parameter) and (see [5] for more details) with

$$\rho_3 = \frac{4.50}{\tau_{cell}} (\mathsf{y}_0 \rho_\star / \rho_1), \text{ and } E(t) = \begin{cases} \exp(\dfrac{1}{t^2 - 1}) \text{for } |t| < 1, \\ 0 \text{ for } |t| \geq 1 \end{cases}, \tag{2.1b}$$

$$\psi(t) \equiv \psi(\rho_3; t) = \rho_3 E(\frac{2t}{\tau} + 1), \text{ for } t \in [-\tau, 0) \tag{2.1c}$$

where τ is synonymous with τ_{cell}. The choice of the scaling factor 4.50 in (2.1b) ensures that $\int_0^\tau \psi(s - \tau_{cell}) ds = \mathsf{y}_0$. Thus, $\psi(t)$ depends upon parameters to be assigned or to be estimated. In the linear NDDE model under consideration, one possible biological interpretation of the parameters is given in Table 1; the solution will be denoted

$$y(t) \equiv y(\mathbf{p}; t) = y([\rho_{0,1,2,3}, \tau]; t), \tag{2.1d}$$

where [vi] $\mathbf{p} = [\rho_{0,1,2,3}, \tau]$. In the simplest scenario, we suppose that we have N obser-

Notation	Biological interpretation	Units
$\tau > 0$;	the average cell-division time	min
$-\rho_0 \geq 0$	the rate of cell-death	min^{-1}
$\rho_1 \geq 0$	the rate of commitment to the cell-division process	min^{-1}
$0 \leq \rho_2 \leq 2$	the gradual dispersal of synchronization of cell-division ($\rho_2 = 2$ implies perfect synchronization)	
$0 \leq \rho_\star \leq 1$	the fraction of cells dividing over the first step	

Table 1: A biological interpretation of the parameters in the model (2.1).

vations pairs, $\{t_j; \mathsf{y}_j\}_{j=1}^N$ with

$$0 = t_0 \leq t_1 < t_2 < \cdots < t_{N-1} < t_N, \text{ where } \mathsf{y}_0 = \mathsf{y}_1 \text{ is known if } t_0 = t_1,$$

[vi] Or $\mathbf{p} = [\rho_{0,1,2,3}, \tau, \mathsf{y}_0]$ if y_0 is unknown and has to be estimated.

and we seek a parameter $\mathbf{p}$ in our computational model such that $y(\mathbf{p};\mathsf{t}_j) \approx \mathsf{y}_j$. The model-fitting problem is to select that parameter $\mathbf{p}_*$ for which the function $\mathbf{y}(t;\mathbf{p}_*)$ provides a 'best' fit, at arguments $t = \mathsf{t}_j$, to the given data set $\{\mathsf{y}_j\}_{j=1}^N$.

2.4 Interpretation of the model equations

Before proceeding too far, we should take stock (given the application envisaged) of the interpretation of the equations that define our model. We shall suppose that we are attempting to model a process in which the cell population grows continuously with time t. This has a bearing on the definition of the initial function ψ, which suffers a jump discontinuity at the right end-point $t = t_0$ (and has no left-hand derivative at this point). It is necessary to observe that $\psi(t)$ does not represent the state of the population for $t \in [t_0 - \tau, t_0]$ but is an "initial function" chosen to ensure that the model is reasonable for $t \geq t_0$; compare the observation in footnote iv. However, if the model equations are inappropriately interpreted, the solution can appear to suffer jump discontinuities.

Let us first consider a solution in the sense of Carathéodory; that is, we seek an absolutely continuous function that satisfies the integrated form of (1.3), *viz*

$$y(t) - \mathsf{y}_0 = \int_{t_0}^{t} \{\rho_0 y(s) + \rho_1 y(s-\tau) + \rho_2 y'(s-\tau)\} \, \mathrm{d}s. \qquad (2.2)$$

For $t \in [t_0, t_0 + \tau]$ this yields

$$y(t) - \int_{t_0}^{t} \rho_0 y(s) \, \mathrm{d}s = \mathsf{y}_0 + \int_{t_0-\tau}^{t-\tau} \{\rho_1 \psi(s) + \rho_2 \psi'(s)\} \, \mathrm{d}s \qquad (2.3)$$

(and it will be noted that the value, at $t = t_0$, of the integrands in the second integral does not affect the value of this integral). The last equation forms the foundation of a method of steps for solving the equation: we consider, for $n \in \{0, 1, 2, \cdots\}$ the equation

$$y(t) - \int_{t_n}^{t} \rho_0 y(s) \, \mathrm{d}s = y(t_n) + \int_{t_n-\tau}^{t-\tau} \{\rho_1 y(s) + \rho_2 y'(s)\} \, \mathrm{d}s \text{ where } t \in [t_n, t_{n+1}], t_n = n\tau.$$

For the method of steps, it only remains to note that the solution of $y(t) - \int_{t_n}^{t} \rho_0 y(s) \, \mathrm{d}s$

$= g_n(t)$ $(t_n \leq t \leq t_{n+1})$ is $y(t) = \exp\{\rho_0(t - t_n)\}g_n(t_n) + \int_{t_n}^{t} \exp\{\rho_0(t - s)\}g_n'(s) \, \mathrm{d}s$
for $t_n \leq t \leq t_{n+1}$ and we can use this result for the previously displayed equation, with $g_n'(s) = \rho_1 y(s - \tau) + \rho_2 y'(s - \tau)$. The solution of the integrated form is also a continuous solution of (1.3a) or, equally, a continuous solution of (1.3b) when derivatives are (re)interpreted as right-hand derivatives.

THEOREM 2.1 *The solution in the sense of Carathéodory of our model equations is*

also a solution of $L_+\{y\}(t) = 0$ for $t \in [t_0, T)$. Thus,

$$\{y(t) - \rho_2 y(t - \tau)\}'_+ = \rho_0 y(t) + \rho_1 y(t - \tau) \quad (t \in [t_0, T)) \tag{2.4a}$$

and

$$y'_+(t) = \rho_0 y(t) + \rho_1 y(t - \tau) + \rho_2 y'_+(t - \tau) \quad (t \in [t_0, T)) \tag{2.4b}$$

with

$$y(t) = \psi(t), \quad (and \; y'_+(t) = \psi'_+(t)) \quad for \; t \in [t_0 - \tau, t_0]. \tag{2.4c}$$

2.5 Reformulating our cell-growth model

We return to the combination of the NDDE and the initial function that combine to form our cell growth model (2.1). Related to the solution that satisfies (2.1) are two functions $y_0(t)$ and $y_1(t)$ each satisfying the NDDE (2.1a) but with differing (complementary) initial conditions. The following superposition principle holds:

THEOREM 2.2 The solution $y(t) \equiv y([\rho_{0,1,2,3}, \tau]; t)$ of (2.1) satisfies

$$y(t) = y_0(t) + y_1(t). \tag{2.5a}$$

where $y_0(t) \equiv y_0([\rho_{0,1,2,3}, \tau]; t)$, $y_1(t) \equiv y_0([\rho_{0,1,2}, \tau]; t)$ satisfy (2.1a), and where

$$y_0(t) = \rho_3 E(2t/\tau + 1), \quad t \in [-\tau, 0], \tag{2.5b}$$

$$y_1(t) = 0, \quad t \in [-\tau, 0), \quad y_1(0) = \mathsf{y}_0 \tag{2.5c}$$

and, in the notation of (1.7a),

$$L_+\{y_0\}(t) = 0 \; and \; L_+\{y_1\}(t) = 0 \; for \; t \in [0, T) \tag{2.5d}$$

It is useful to have the notation $y_\sharp$ for the solution of

$$L_+\{y_\sharp\}(t) = 0 \; for \; t \in [0, T), \quad y_\sharp(t) = 0, \quad t \in [-\tau, 0), \quad y_\sharp(0) = 1, \tag{2.6a}$$

since in this notation $y_1(t) = \mathsf{y}_0 y_\sharp(t)$. Clearly, $y'_\sharp(t) = \rho_0 y_\sharp(t)$, for $t \in [0, \tau]$ and we have

$$y'_\sharp(t) = \rho_0 y_\sharp(t) + \rho_1 y_\sharp(t - \tau) + \rho_1 y'_\sharp(t - \tau) \; for \; t \geq \tau, \tag{2.6b}$$

$$y_\sharp(t) = \exp(\rho_0 t) \; on \; [0, \tau]. \tag{2.6c}$$

Remark 2.1 (a) Observe that $y_0 \in C^1[-\tau, 0]$ while y_1 (which is independent of ρ_3) has the jump discontinuity but is continuous for $t \geq 0$. (b) According to Remark 1.1, one can normalize $\tau > 0$ to have the value unity and consider $u(t) = y(\tau t) = y_0(\tau t) + y_1(\tau t)$.

10

2.6   Parameter selection

The key part in selecting values of parameters in order to fit a model to data is the formulation of an objective function to be optimized by the choice of these values. We seek the parameter $\mathbf{p}_\star$ for which the corresponding values $\{y(\mathbf{p}_\star; \mathbf{t}_j)\}_{j=1}^N$, provide a 'best fit' to the given data $\{\mathbf{y}_j\}_{j=1}^N$ in the sense that

$$\Phi(\mathbf{p}_\star) = \min_{\mathbf{p}} \Phi(\mathbf{p}). \tag{2.7}$$

Different fitness functions can be used [17], depending on the statistical features of the errors in the data. There is a variety of methods for regression analysis and interpretation of statistical properties of estimation schemes [16–18,25,36]. For a comparison of objective functions, we refer to [43].

A general framework for consistent parameter estimation is provided by the maximum likelihood approach. In addition, this approach establishes a common basis for ranking models that may vary considerably in their "architecture". In this framework the data is regarded as fixed and the parameters as variable. Those parameters for which the likelihood of obtaining exactly the observed data is the highest:

$$\mathcal{L}(\mathbf{p}_\star) = \max_{\mathbf{p}} \mathcal{L}(\mathbf{p}), \tag{2.8}$$

represent the "maximum likelihood estimates". The statistical properties of the errors in the observed data are assumed to be expressible using an appropriate probability density distribution (*e.g.*, a normal, or log-normal, distribution).

THEOREM 2.3 *(Likelihood function) If the errors in the observed date at successive times are independent and have a Gaussian (normal) distribution about the vectors $y_j := \{y(\mathbf{p}; \mathbf{t}_j)\}_{j=1}^N$, that is, $\mathbf{y}_j \sim N(y_j, \sigma_j)$, where σ_j is the j-th covariance matrix, then the likelihood function can be written as*

$$\mathcal{L}(\mathbf{p}) = \prod_{j=1}^N \mathcal{H}(\mathbf{y}_j; \mathbf{p}). \tag{2.9}$$

Here, the component probability density functions are given by

$$\left\{ \mathcal{H}(y_j; \mathbf{p}) = \frac{1}{\sqrt{(2\pi)\sigma_j^2}} \exp\{-\frac{1}{2}[y_j - \mathbf{y}_j]^T \sigma_j^{-2}[y_j - \mathbf{y}_j]\} \right\}. \tag{2.10}$$

Under the further assumption that the data variance is equal to the corresponding maximum likelihood estimate, the *maximum likelihood* approach is equivalent to the common *weighted least squares* parameter estimation. The discussion for the present example will be based, in the main, on the use of the weighted objective function:

$$\Phi(\mathbf{p}) = \sum_{j=1}^{N} w_j \left[y(\mathbf{p}; \mathsf{t}_j) - \mathsf{y}_j \right]^2, \text{ with the weights } w_j > 0 \text{ defined by } w_j = \sigma^{-2}\omega_j,$$

(2.11)

as the measure of the discrepancy between the model forecast and the observed data.

Remark 2.2 *As part of a more general discussion, it is useful to consider alternative objective functions. To assist such discussions we can denote (2.11) with constant weights unity by* $\Phi_{\scriptscriptstyle OLS}$:

$$\Phi_{\scriptscriptstyle OLS}(\mathbf{p}) = \sum_{j=1}^{N} \left[y(\mathbf{p}; \mathsf{t}_j) - \mathsf{y}_j \right]^2;$$

(2.12a)

with arbitrary non-negative weights $\{w_j\}$ *by* $\Phi_{\scriptscriptstyle WLS}$,

$$\Phi_{\scriptscriptstyle WLS}(\mathbf{p}) = \sum_{j=1}^{N} w_j \left[y(\mathbf{p}; \mathsf{t}_j) - \mathsf{y}_j \right]^2 \text{(for any non-trivial choice with } w_j \geq 0).$$

(2.12b)

The similar (but different) objective function associated with a log-least-squares approach is

$$\Phi_{\scriptscriptstyle LLS}(\mathbf{p}) = \sum_{j=1}^{N} \left[\ell n \left(y(\mathbf{p}; \mathsf{t}_j) \right) - \ell n \left(\mathsf{y}_j \right) \right]^2$$

(2.12c)

(the data and solution values must be positive). Minimization of (2.12a), (2.12b), (2.12c) defines, respectively, the "OLS"-approach, "WLS"-approach, and the "LLS"-approach.

To solve (2.8) we seek $\mathbf{p}_{\star} = \arg\min \Phi(\mathbf{p})$.

THEOREM 2.4 *Given sufficient smoothness,*

$$\Phi(\mathbf{p}+\varepsilon\boldsymbol{\delta}\mathbf{p}) = \Phi(\mathbf{p}) + 2\sum_{j=1}^{N} w_j \Big\{ y(\mathbf{p}; \mathsf{t}_j) - \mathsf{y}_j \Big\} \Big\{ y(\mathbf{p}+\varepsilon\boldsymbol{\delta}\mathbf{p}; \mathsf{t}_j) - y(\mathbf{p}; \mathsf{t}_j) \Big\} + \mathcal{O}(\varepsilon^2),$$ (2.13)

(irrespective of the choice of positive weights), and, provided that the derivatives exist,

$$\frac{\partial}{\partial p_\ell}\Phi(\mathbf{p}) = 2\sum_{j=1}^{N} w_j \left[y(\mathbf{p}; \mathsf{t}_j) - \mathsf{y}_j \right] \frac{\partial}{\partial p_\ell} y(\mathbf{p}; \mathsf{t}_j).$$

(2.14)

Parameter values for which (2.14) vanishes are candidates[vii] for the choice as op-

---

[vii] Since τ is constrained to be positive, the possibility that an optimum parameter occurs with $\tau = 0$ should be admitted; a similar observation applies with other parameters that are meaningful only if constrained (see Table 1). Note that the system of equations $\frac{\partial}{\partial p_\ell}\Phi(\mathbf{p}) = 0$ for $l \in \{1, 2, \cdots, L\}$ (or the similar equations using right-hand derivatives) may be ill-conditioned.

timal parameter. Consequently, Theorem 2.4 demonstrates a practical rôle for the values $\frac{\partial}{\partial p_\ell} y(\mathbf{p}; \mathsf{t}_j)$ of the *first-order sensitivity coefficients* evaluated at the data-sampling abscissae. When Φ is piecewise-smooth we may find it necessary to replace derivatives in (2.14) by right-hand derivatives. In fact, it was pointed out by Baker and Paul [11] that with certain choices of parameters the function Φ may display jump discontinuities, and standard optimization routines may require adaptation or monitoring.

2.7 Selection from a hierarchy of models

Model identification may be regarded as having two components: (*i*) computation of the values of the parameters by fitting observational data set and (*ii*) selection, from a range of forms, of a parameterized mathematical model. Choosing any one type of model (ODE, DDE, NDDE) from amongst those considered here corresponds to an *a priori* restriction of the number of parameters. Thus, if we denote the corresponding formal parameter by $\mathbf{p}^{ODE}$, $\mathbf{p}^{DDE}$, and $\mathbf{p}^{NDDE}$, the corresponding vector spaces $\{V_*\}$ of parameters are nested: $V^{ODE} \subset V^{DDE} \subset V^{NDDE}$. In this sense, the ODEs, DDEs, and NDDEs considered here form a nested hierarchy of models. The parameter selection procedure will depend upon a choice of an objective function $\Phi(\mathbf{p})$ in (2.11) and the determination of $\mathbf{p}_\star^{ODE}$, $\mathbf{p}_\star^{DDE}$, and $\mathbf{p}_\star^{NDDE}$ such that

$$\mathbf{p}_\star^{ODE} = \arg \min_{\mathbf{p}^{ODE} \in V^{ODE}} \Phi(\mathbf{p}), \quad \mathbf{p}_\star^{DDE} = \arg \min_{\mathbf{p}^{DDE} \in V^{DDE}} \Phi(\mathbf{p}), \quad \mathbf{p}_\star^{NDDE} = \arg \min_{\mathbf{p}^{NDDE} \in V^{NDDE}} \Phi(\mathbf{p}). \tag{2.15}$$

Clearly, because the models form a hierarchy $\Phi(\mathbf{p}_\star^{ODE}) \leq \Phi(\mathbf{p}_\star^{DDE}) \leq \Phi(\mathbf{p}_\star^{NDDE})$, but we now proceed with one of the standard selection principles to decide on the most appropriate model.

It is clear that an increase in the number of model parameters reduces bias in the data fit. However, small data sets do not support models with many parameters since the increase (above some critical value) in the number of parameters to be estimated from a finite data set leads to an increase of the variance in the parameters estimates [viii] . Therefore, the complexity of the model structure has to be justified. There are various theoretical bases (Bayesian, information-theoretic, or minimum description length principles) for making such a selection, but a common feature of many of them is that the models are ranked in order of some multiple of $\ell n\, \Phi(\mathbf{p})$ (data description bias) modified by the addition of an *regulator* or *penalty measure* μ, to yield (as, *e*.g., in (2.19a) below) an index

$$\kappa(\mathbf{p}) := N\ell n\, \Phi(\mathbf{p}) + \mu \tag{2.16}$$

[viii]Notice that a quantitative characterization of the amount of information (about the model parameters) that observable variables carry is provided by the Cramer-Rao information inequality [ix] [44,46]. This requires the computation of the Fisher Information Matrix [24, p. 358] defined via the expectation of the second derivative of the log of the likelihood function with respect to parameters.

where μ may depend on the dimensionality of $\mathbf{p}$, the number of data points, or information obtained from (say) the Fisher Information Matrix. The bias term $\ell n\,\Phi(\mathbf{p}_\star)$ indicates the link of the selection criterion with the maximized likelihood function through the relation

$$-2\ell n\,\mathcal{L}(\mathbf{p}) = N\ell n(2\pi) + \sum_{j=1}^{N} \ell n(w_j) + \ell n\Phi(\mathbf{p}) \tag{2.17}$$

Remark 2.3 *The Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC) are widely used as criteria of model selection tools. If N is the number of data points, L is the number of estimated parameters, and $\Phi(\mathbf{p})$ is the* weighted sum *of squared residuals with uncorrelated Gaussian errors characterized by the coefficient of variation σ^2, with*

$$\Phi(\mathbf{p}) = \sigma^{-2}\Phi_\omega(\mathbf{p}) := \sigma^{-2} \sum_{j=1}^{N} \omega_j \left[y(\mathbf{p};\mathsf{t}_j) - \mathsf{y}_j \right]^2, \tag{2.18}$$

then

$$\kappa_{AIC} = N\ell n\,\Phi_\omega(\mathbf{p}_*) + 2(L+1) \quad (\mu = 2(L+1)), \tag{2.19a}$$
$$\kappa_{SIC} = N\ell n\,\Phi_\omega(\mathbf{p}_*) + (L+1)\ell n\,N \quad (\mu = (L+1)\ell n\,N). \tag{2.19b}$$

The best model is then defined that which yields the lowest value of κ_{AIC} or κ_{SIC}.

Remark 2.4 *With the notation in Remark 2.2, we can define the notation $\kappa_{AIC}^{OLS}, \kappa_{SIC}^{OLS}, \kappa_{AIC}^{WLS}, \kappa_{SIC}^{WLS}, \kappa_{AIC}^{LLS}, \kappa_{SIC}^{LLS}$ by replacing $\Phi_\omega(\mathbf{p}_*)$ in (2.19a) by, respectively, $\Phi_{OLS}(\mathbf{p}_*), \Phi_{WLS}(\mathbf{p}_*)$, or $\Phi_{LLS}(\mathbf{p}_*)$. Because we are interested in the relative size of the indicators for different models, extraneous terms associated with the different types of objective functions can be ignored.* The conditions that motivate and justify the use of each of the indices differ from each other.

We estimated the parameters for the set of ODE-, DDE-, and NDDE models under the assumption of identical weights $\omega_j = 1, j = 1, \ldots, N$, i.e. equal data variance σ^2 at successive observation times. The parameter estimation results together with the computed values of the above information-theoretic criteria are summarized in Table 2. They clearly suggest that the NDDE model provides the most coherent description of the synchronized cell population growth data.

In completion, it should be noticed that the model's complexity measure in the information-theoretic framework is incomplete. It has been proposed that the minimum description length (MDL) criterion takes into account the goodness of the fit to the data and the structural and functional complexity of the model in a more comprehensive way. It prescribes that the model with the smallest value of the score
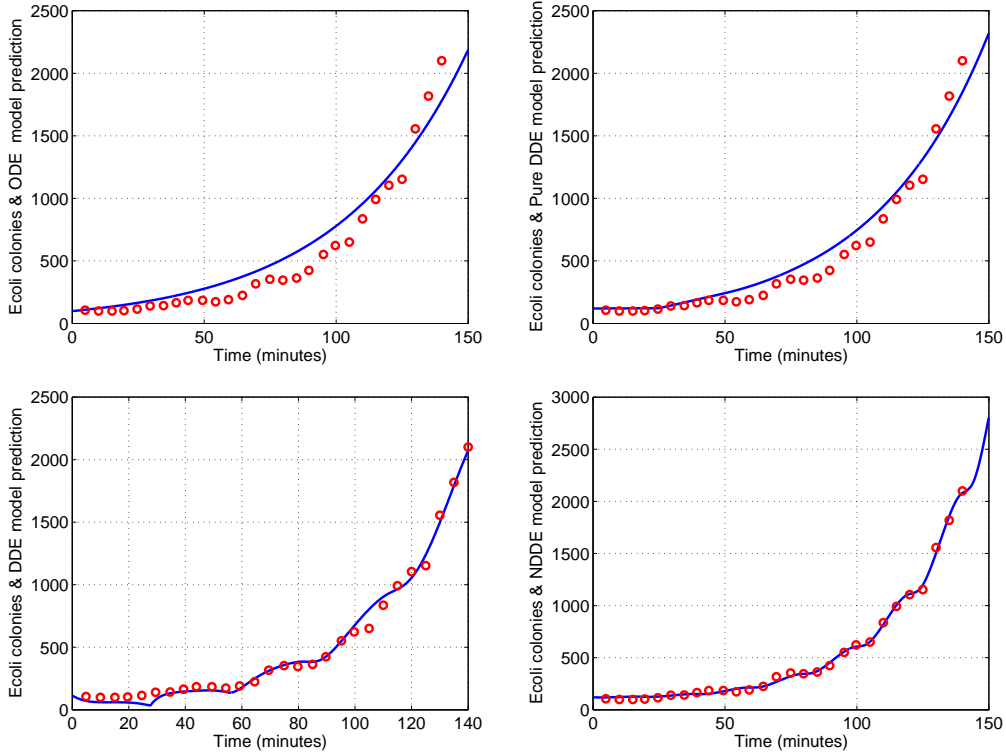
Fig. 2. The top-left panel shows prediction of the exponential growth model (2.21a) (solid line) to the data of *E.coli* colonies (circles). The top-right panel shows the solution of the pure DDE (2.21b) fitted to the *E.coli* data. The bottom-left panel shows the solution of the DDE (2.21c) fitted to the *E.coli* data. The bottom-right panel shows the consistency of the NDDE model (2.21d) with the *E.coli* data.

should be regarded to be superior over the others in the set as it permits the greatest compression of the data in its description. The MDL framework requires the computation of the following score:

$$\kappa_{MDL} = -\ell n \mathcal{L}(\mathbf{p}_\star) + \frac{(L+1)}{2} \log(\frac{N}{2\pi}) + \log(\int_V \sqrt{\det[\mathbf{F}(\mathbf{p})]} d\mathbf{p}), \qquad (2.20)$$

where V is the parameter vector space and $\mathbf{F}(\cdot)$ stands for the Fisher information matrix. We refer for further discussion to [29].

EXAMPLE 2.1 *The modelling problems that we consider are as follows:*

(a) *Estimate* $\tau_{calture} = \dfrac{\ell n\, 2}{\rho_0}$ *in the simple exponential growth model (with initial value* $y(0) = \mathsf{y}_0$*)*

15

$$y'(t) = \rho_0 y(t) \quad (t \geq 0); \tag{2.21a}$$

(b) *Obtain an estimate of* $\tau = \tau_{cell} = \dfrac{1}{\rho_1}$, *with* $\rho_0 = \rho_2 = 0$, *and* $\rho_\star = 1$ *(with given initial function* $\psi(t)$ *described in (2.1c)) in the pure time-lag growth model*

$$y'(t) = \rho_1 y(t - \tau) \quad (t \geq 0); \tag{2.21b}$$

(c) *Obtain an estimate of* $\tau = \tau_{cell}$, ρ_0, *and* ρ_1, *with* $\rho_2 = 0$, *and* $\rho_\star = 1$ *(with given initial function* $\psi(t)$ *described in (2.1c)), in the model*

$$y'(t) = \rho_0 y(t) + \rho_1 y(t - \tau) \quad (t \geq 0); \tag{2.21c}$$

(d) *Obtain an estimate of* $\tau = \tau_{cell}$, ρ_0, ρ_1, ρ_2, $\rho_\star$ *in (2.1c) and in*

$$y'(t) = \rho_0 y(t) + \rho_1 y(t - \tau) + \rho_2 y'(t - \tau) \quad (t \geq 0); \tag{2.21d}$$

The solutions of the models (2.21a)-(2.21d) corresponding to the optimal parameters appear in Figure 2. Numerical results for the parameter estimates and κ_{AIC} *&* κ_{SIC} *(defined by (2.19)) are displayed in Table 2.6a. There is a noticeable qualitative and quantitative improvement on moving through the hierarchy of models (2.21b)–(2.21d).*

2.8   Sensitivity of model predictions to parameter estimates

As mentioned in the introduction of this paper, the information contained in the sensitivity analysis is useful for parameter identification, optimization, reduction of complex nonlinear models, and for experimental design and analysis. Of considerable importance in assessing the acceptability of a model (2.1), is the sensitivity of the model solution $y(\mathbf{p}; t)$ to small variations in the parameter $\mathbf{p}$. For example, if it can be observed that a particular parameter p_j has little effect on the solution, it may be possible to eliminate it, at some stage, from the modelling process. (Arguably, ranking within a hierarchy models using a suitable modelling index would also determine this course.)

3   Perturbation theory

We give varying approaches to the analysis of sensitivity for models described by *neutral delay differential equations* (NDDEs).

3.1   Dependency of the solution on the model

We provide a result that demonstrates the dependence of a solution of the inhomogeneous form of (1.3) on the parameters, initial function, and inhomogeneous term.

Model	ρ_0	ρ_1	ρ_2	ρ_3	τ	$\|Err^{OLS}\|_2$ (no. iters.)	κ^{OLS}_{AIC} & κ^{OLS}_{SIC} (2.19)
ODE (2.21a)		-	-	-	33.56	724 (6)	372.75 & 375.41
DDE (2.21b)		-	-	-	26.220	571.80 (6)	359.53 & 362.20
DDE (2.21c)	-0.1165	0.3088	-	-	27.716	258.89 (15)	319.16 & 324.49
NDDE (2.21d)	-0.0057	0.0131	1.8407	0.1600	20.2229	128.37 (13)	283.87 & 291.87

Table 2.6a: Parameter estimates, errors (with the number of iterations [39]), and κ^{OLS}_{AIC} & κ^{OLS}_{SIC} values for *E.coli* growth models, using the OLS approach (see Remark 2.2); here, $w_i = 1$ and $\|Err^{OLS}\|_2 = \sqrt{\Phi_{OLS}(\mathbf{p}_*)}$.

Model	ρ_0	ρ_1	ρ_2	ρ_3	τ	$\|Err^{WLS}\|_2$ (no. iters.)	κ^{WLS}_{AIC} & κ^{WLS}_{SIC} (2.19)
DDE (2.21b)		-	-	-	28.820	2.506 (4)	55.46 & 58.13
DDE (2.21c)	-0.1120	0.3032	-	-	27.650	1.1679 (12)	16.69 & 22.02
NDDE (2.21d)	-0.0059	0.0131	1.844	0.1794	20.216	0.5073 (15)	-26.01 & -18.03

Table 2.6b: Parameter estimates, errors (with the number of iterations), and κ^{WLS}_{AIC} & κ^{WLS}_{SIC} values for *E.coli* growth models, using the WLS approach (see Remark 2.2); here, $w_i = \frac{1}{Y_i^2}$ and $\|Err^{WLS}\|_2 = \sqrt{\Phi_{WLS}(\mathbf{p}_*)}$.

Model	ρ_0	ρ_1	ρ_2	ρ_3	τ	$\|Err^{LLS}\|_2$ (no. iters.)	κ^{LLS}_{AIC} & κ^{LLS}_{SIC} (2.19)
DDE (2.21c)		-	-	-	30.776	1.346 (5)	20.68 & 20.34
DDE (2.21c)	-0.0603	0.1565	-	-	26.462	0.7562 (10)	-7.65 & -2.32
NDDE (2.21d)	-0.0072	0.0183	1.6654	0.1704	20.172	0.4852 (4)	-28.49 & -20.50

Table 2.6c: Parameter estimates, errors (with the number of iterations), and κ^{LLS}_{AIC} & κ^{LLS}_{SIC} values for *E.coli* growth models, using the LLS approach; see Remark 2.2. Here, $\|Err^{LLS}\|_2 = \sqrt{\Phi_{LLS}(\mathbf{p}_*)}$.

THEOREM 3.1 *(See Baker & Parmuzin [14]) There exist functions $\mathcal{C}$, $\mathcal{K}$, and $\mathcal{Y}$ (collectively dependent on $\mathbf{p} = [\rho_{0,1,2}, \tau]$) such that the solution $y(t) = y(\mathbf{p}, \psi; t)$ of*

$$\{y(t) - \rho_2 y(t - \tau)\}' = \rho_0 y(t) + \rho_1 y(t - \tau) + f(t) \quad (t \in [0, T)) \qquad (3.1)$$

with

$$y(t) = \psi(t) \ (t \in [-\tau, 0]) \qquad (3.2)$$

(where $\psi \in C(-\tau, 0)$) can be written in the form

$$y(\psi; t) = \left\{ \mathcal{Y}(0, t)(\mathfrak{U}\psi)(0) + \mathcal{C}(t)\psi(t - \lfloor \frac{t}{\tau} + 1 \rfloor \tau) + \int_{-\tau}^{0} \mathcal{K}(s, t)\psi(s)ds \right\} + \qquad (3.3)$$

$$+ \int_{0}^{t} \mathcal{Y}(s, t)f(s)ds, \ \text{wherein} \ \mathfrak{U}\psi(0) := \psi(0) - \rho_2 \psi(-\tau).$$

A constructive proof of this result, which can be regarded as a *variation of parameters* result, yields a means for generating expressions for $\mathcal{C}$, $\mathcal{K}$, and $\mathcal{Y}$. A feature

17

that is important to us is embodied in the following consequence of Theorem 3.1.

THEOREM 3.2 *With* $\psi \in C(-\tau, 0)$ *in (3.2),* $y(t)$ *depends, at any* $t \geq 0$*, on values* $\psi(s)$ *for* s *ranging over* $[-\tau, 0]$.

The preceding results extend in an obvious manner to the case where $t_0 \neq 0$ (see Remark 1.1).

3.2 Sensitivity

From the foregoing discussion, a key feature in the determination of a good model is *the response of a solution of the model to changes in the model*. This includes the response to changes in the initial function ψ. The existence theory reveals that the continuity class of $y(\mathbf{p} + \varepsilon\boldsymbol{\delta}\mathbf{p}, \psi + \delta\psi; \cdot)$ may differ from that of $y(\mathbf{p}, \psi; \cdot)$. Where they exist, the (possibly one-sided) limits

$$\lim_{\varepsilon \to 0+} \frac{y(\mathbf{p} + \varepsilon\boldsymbol{\delta}\mathbf{p}, \psi + \varepsilon\delta\psi; \cdot) - y(\mathbf{p}, \psi; \cdot)}{\varepsilon}, \tag{3.4a}$$

define various sensitivities any one of which is obtained for specific

$$\boldsymbol{\delta}\mathbf{p} \text{ with } \|\boldsymbol{\delta}\mathbf{p}\| = 1 \text{ and/or } \delta\psi \text{ with } \|\delta\psi\| = 1 \tag{3.4b}$$

(here, $\delta\psi$ may vanish, or selected components δp_ℓ may vanish). These have a significant role in determining parameters that are in some sense optimal given observed data. They also have a rôle (i) in determining qualitative behaviour of a solution, and (ii) in making a choice of numerical method for computing a solution. As a special case, the limits corresponding to variations in individual components of $\mathbf{p}$ give rise to

$$\frac{\partial y}{\partial p_\ell}(\mathbf{p}, \psi; t) \quad (t \in [t_0, T), \tag{3.5}$$

which (when the derivative exists) is a first-order *sensitivity coefficient*.

If $\delta\mathbf{p} = 0$ and one considers the limit (3.4a) for a fixed and normalized function $\delta\psi$ then one is, in effect, invoking a Gateau derivative to define the limit. However, where ψ is defined by one or more parameters $\{p_\lambda\}_{\lambda \in \Lambda}$ one can investigate (3.5) for $\ell \in \Lambda$.

3.3 Finite difference simulation of the first-order sensitivity

To illustrate, we consider the solution $y(t)$ of the our cell growth model equations, and we regard $\rho_0, \rho_1, \rho_2, \tau$ and y_0 as parameters. Thus, $y(t) \equiv y([\rho_0, \rho_1, \rho_2, \tau, \mathsf{y}_0]; t)$. Perturbations in the parameters give rise to corresponding perturbations in the solution, *e.g.*,

$$\delta_0 y(t) := \{ y([\rho_0 + \delta\rho_0, \rho_1, \rho_2, \tau, \mathsf{y}_0]; t) - y([\rho_0, \rho_1, \rho_2, \tau, \mathsf{y}_0]; t) \} \tag{3.6a}$$

$$\delta_1 y(t) := \{ y([\rho_0, \rho_1 + \delta\rho_1, \rho_2, \tau, \mathsf{y}_0]; t) - y([\rho_0, \rho_1, \rho_2, \tau, \mathsf{y}_0]; t) \} \tag{3.6b}$$

$$\delta_2 y(t) := \{ y([\rho_0, \rho_1, \rho_2 + \delta\rho_2, \tau, \mathsf{y}_0]; t) - y([\rho_0, \rho_1, \rho_2, \tau, \mathsf{y}_0]; t) \} \tag{3.6c}$$

$$\delta_{\mathsf{y}_0} y(t) := \{ y([\rho_0, \rho_1, \rho_2, \tau, \mathsf{y}_0 + \delta\mathsf{y}_0]; t) - y([\rho_0, \rho_1, \rho_2, \tau, \mathsf{y}_0]; t) \} \tag{3.6d}$$

$$\delta_\tau y(t) := \{ y([\rho_0, \rho_1, \rho_2, \tau + \delta\tau, \mathsf{y}_0]; t) - y([\rho_0, \rho_1, \rho_2, \tau, \mathsf{y}_0]; t) \}. \tag{3.6e}$$

The perturbations clearly respond to the size of the change in the parameters. The *first-order sensitivity functions* (3.5) can be approximated by finite-difference techniques. Then the first-order sensitivity functions can be approximated by computing, respectively,

$$\frac{\delta_0 y(t)}{\delta\rho_0} \approx (\frac{\partial}{\partial\rho_0})_+ y(t), \quad \frac{\delta_1 y(t)}{\delta\rho_1} \approx (\frac{\partial}{\partial\rho_1})_+ y(t), \quad \frac{\delta_2 y(t)}{\delta\rho_2} \approx (\frac{\partial}{\partial\rho_2})_+ y(t), \tag{3.7a}$$

$$\frac{\delta_2 y(t)}{\delta\tau} \approx (\frac{\partial}{\partial\tau})_+ y(t) \tag{3.7b}$$

and (see below)

$$\frac{\delta_{\mathsf{y}_0} y(t)}{\delta\mathsf{y}_0} := \{ y([\rho_0, \rho_1, \rho_2, \tau, \mathsf{y}_0 + \delta\mathsf{y}_0]; t) - y([\rho_0, \rho_1, \rho_2, \tau, \mathsf{y}_0]; t) \} / \delta\mathsf{y}_0 \tag{3.7c}$$

where $\delta\rho_{0,1,2} > 0$ $\delta\tau > 0$ and $\delta\mathsf{y}_0 > 0$. Every function in (3.7a) satisfies the same initial condition.

In the case of (3.7b) the change in τ induces a change in the initial[x] condition (2.1c) which (applied to $y([\rho_0, \rho_1, \rho_2, \tau + \delta\tau, \mathsf{y}_0]; t)$) now becomes

$$\psi(t) = \rho_3 E(\frac{2t}{\tau + \delta\tau} + 1), \text{ for } t \in [-\tau - \delta\tau, 0), \psi(0) = \mathsf{y}_0. \tag{3.8}$$

Turning to sensitivity with respect to y_0, it is easy to see that the first-order sensitivity function (and its difference simulation) are both expressible in the notation (2.6a) as

$$\frac{\partial}{\partial\mathsf{y}_0} y(t) = y_\sharp(t). \tag{3.9}$$

Figure 3 shows the functions[xi] in (3.7) corresponding to small changes to the parameters $\mathsf{y}_0, \rho_0, \rho_1, \rho_2,$ and τ (changes $\delta y_0 = 5$; $\delta\rho_0 = 0.01$; $\delta\rho_1 = 0.01$; $\delta\rho_2 = 0.1$; $\delta\tau = 2$). These functions can be used to indicate which parameters have a significant effect on the solution; they demonstrate the measure of the importance of the input parameters, and enable one to assess the relevant time intervals for the identification of specific parameters and enhance the understanding of the role played by specific

---

[x] If one seeks sensitivity with respect to ρ_3 in (2.1c), a change in the initial function also occurs.

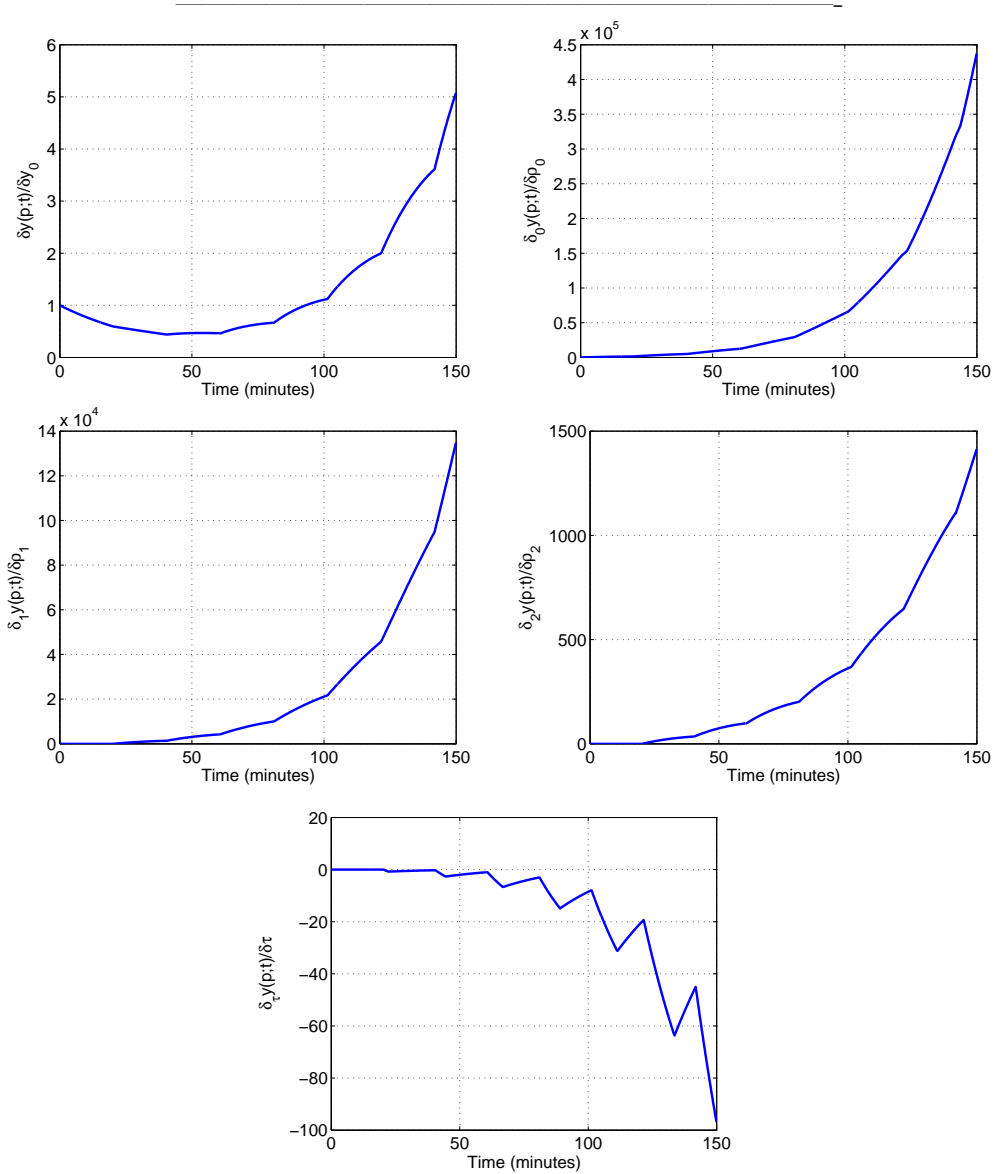[xi] Parameterized solutions were computed using `Archi` [39] with tolerance 10^{-9}.

Fig. 3. The figures show a selection of finite difference simulations in (3.7) due to changes in the parameters [y_0 (top-left), ρ_0 (top-right), ρ_1 (middle-left), ρ_2 (middle-right), and τ (bottom)] The plots have "kinks" at $t = n\tau$, $n = 1, 2, 3, \ldots$ as result of the existence of a delay in the system.

model parameters in describing experimental data. See §3.9. We note from Figure 3 that the solution of the NDDE model is very sensitive to a small change in ρ_0. The oscillation accompanied by $\delta y(t; p)/\delta \tau$ signifies that the solutions of the models (2.21c)&(2.21d) are sensitive to changes in the parameter τ and this parameter plays a significant role in the model. However, robustness is shown with small changes in the initial value y_0.

3.4   An illustration

The estimate of the data variance provided by the maximum likelihood approach, with ω_j identically 1, follows from the optimality condition and is given by the value $\sigma_*^2 = \dfrac{1}{N}\Phi_\omega(\mathbf{p}_\star)$. For the given data set the NDDE model predicts that $\sigma_* \approx 24$. Notice that the cell population grows from 100 to 2100 cells, implying that the predicted relative accuracy of the measurements should also increase by a corresponding factor.

Sensitivity can be determined numerically by selecting $\mathbf{p}$, $\boldsymbol{\delta p}$ and a value $\epsilon > 0$ and computing the difference approximation $\dfrac{1}{\epsilon}\{y(\mathbf{p}+\boldsymbol{\delta p}, \psi + \epsilon\delta\psi; \cdot) - y(\mathbf{p}, \psi; \cdot)\}$ to the limit (3.4a) above. It is also possible to formulate systems of equations for the first-order sensitivity equations.

3.5   Sensitivity for our cell growth model

We commence with (2.1), satisfied by $y(t) \equiv y([\rho_{0,1,2,3}, \tau]; t)$, and take partial derivatives.

LEMMA 3.3 *Where the derivative $\frac{\partial}{\partial \rho_\ell}y'_+(t)$ is continuous, it can be rewritten*

$$\frac{\partial}{\partial \rho_\ell}y'_+(t) = \Big(\frac{\mathrm{d}}{\mathrm{d}t}\Big)_+\frac{\partial}{\partial \rho_\ell}y(t) \tag{3.10}$$

and a similar observation applies (with a similar condition) for partial derivatives with respect to τ:

$$\Big(\frac{\partial}{\partial \tau}\Big)_+ y'_+(t) = \Big(\frac{\mathrm{d}}{\mathrm{d}t}\Big)_+\Big(\frac{\partial}{\partial \tau}\Big)_+ y(t), \tag{3.11a}$$

$$\Big(\frac{\partial}{\partial \tau}\Big)_+ y(t-\tau) = -y'_+(t-\tau), \tag{3.11b}$$

$$\Big(\frac{\partial}{\partial \tau}\Big)_+ y'_+(t-\tau) \equiv \Big(\frac{\mathrm{d}}{\mathrm{d}t}\Big)_+\Big(\frac{\partial}{\partial \tau}\Big)_+ y(t-\tau) = -y''_+(t-\tau). \tag{3.11c}$$

To show (3.11), $\big\{u\big(t - (\tau + |\delta|)\big) - u(t)\big\}/|\delta| = -\big\{u(t) - u\big((t-|\delta|) - \tau\big)\big\}/|\delta| \to -(\frac{\mathrm{d}}{\mathrm{d}t})_+ u(t-\tau)$ (with $u = y$ or $u = y'_+$, *etc.*) provided $(\frac{\mathrm{d}}{\mathrm{d}t})_+ u(t-\tau)$ is continuous.

We now recall the equations in (2.1) and take partial derivatives. Assuming the conditions of Lemma 3.3, we have (for $t \geq 0$)

$$L_+\big\{y\big\}(t) \quad = 0, \quad L_+\Big\{\frac{\partial}{\partial \rho_0}y\Big\}(t) - y(t) = \quad 0, \tag{3.12}$$

$$L_+\Big\{\frac{\partial}{\partial \rho_1}y\Big\}(t) - y(t-\tau) = 0, \quad L_+\Big\{\frac{\partial}{\partial \rho_2}y\Big\}(t) - y'(t-\tau) = 0, \tag{3.13}$$

$$L_+\Big\{\frac{\partial}{\partial \rho_3}y\Big\}(t) = 0, \tag{3.14}$$

while for $t \in [-\tau, 0]$

$$y(t) = \psi(t), \qquad \frac{\partial}{\partial \rho_0}y(t) = 0, \qquad \frac{\partial}{\partial \rho_1}y(t) = 0, \tag{3.15a}$$

$$\frac{\partial}{\partial \rho_2}y(t) = 0, \; \frac{\partial}{\partial \rho_3}y(t) = \frac{\partial}{\partial \rho_3}\psi(t). \tag{3.15b}$$

Here, $\frac{\partial}{\partial \rho_3}\psi(t)$ has the value

$$\frac{\partial}{\partial \rho_3}\psi(t) = E(2t/\tau + 1) \quad (t \in [-\tau, 0]). \tag{3.16}$$

The equations for sensitivity with respect to τ are more subtle. For further remarks on sensitivity to parameters, including the time lag, of a general nonlinear NDDE system we refer to the earlier work [3]. We can examine sensitivity with respect to τ in the case (1.4d) on the basis that $\frac{\mathrm{d}}{\mathrm{d}\tau}\rho_0 = -1/\tau^2$ $(\tau > 0)$.

3.6 Sensitivity and stability

The term *conditioning* is often used by numerical analysts to indicate the magnitude of the sensitivity of a solution to perturbation in a problem, and the term *stability* is frequently used to describe the propagation of errors in an evolutionary problem (such as a linear recurrence [xii]) when initial values are perturbed.

Remark 3.1 *The study of both conditioning and stability (in many senses, including the latter sense) forms part of a general perturbation theory. The terminology employed by numerical analysts is heavily influenced by the historical development of the subject. The classical emphasis on linear evolutionary equations leads to definitions of stability in terms of qualitative properties of solutions, rather than qualitative properties of* perturbations *to a particular solution. It is perhaps surprising that such definitions still find currency in some examples of contemporary research; the saving grace being that in the* linear *case one can establish rigorously the link between "correctly defined" stability and behaviour of the solutions.*

Three differing concepts of stability (in particular) can be related to each other:

[xii] For perspectives on stability in the context of RFDEs see *e.g.*[33].

(1) *Stability with respect to initial conditions,*
(2) *Stability with respect to persistent perturbations,*
(3) *Stability with respect to structural changes.*

The only example of (3) to be considered here is *stability with respect to parameters (which may include stability with respect to the time-lag*[xiii] *).*

The general notion of stability can be conveyed as follows. Denote by $\mathfrak{T}$ an unbounded set of discrete or continuous values of $t \in [t_0, \infty)$ that comprise the domain of definition of y. Here, $y(t)$ is supposed to be the solution of an evolutionary problem P. Corresponding to this, $\widetilde{y}(t)$ $(t \in \mathfrak{T})$ is supposed to be the solution of a perturbed problem $\widetilde{P}$. We may restrict the type of perturbation in the problem, and we require a suitable measure of the distance, say $\mathfrak{d}(P, \widetilde{P}) = \mathfrak{d}(\widetilde{P}, P)$, between the problem $\widetilde{P}$ and P; then the problems are *neighbouring* if $\mathfrak{d}(P, \widetilde{P})$ is 'small'. We also require a suitable measure, say $\|\widetilde{y}(t) - y(t)\|$, of the change in the solution at the point $t \in \mathfrak{T}$. Then the solution $y(t)$ is locally stable if $\|\widetilde{y}(t) - y(t)\|$ is uniformly bounded for all $t \in \mathfrak{T}$ whenever $\mathfrak{d}(P, \widetilde{P})$ is sufficiently small; it is locally asymptotically stable when it is locally stable and in addition $\lim\limits_{\substack{t \to \infty \\ t \in \mathfrak{T}}} \|\widetilde{y}(t) - y(t)\| = 0$.

The stability (respectively, asymptotic stability) is called *global*, rather than local, if, in these definitions, the size of the term $\mathfrak{d}(P, \widetilde{P})$ is unrestricted.

3.7 Types of perturbation

The types of stability enumerated (1)–(3) above correspond to differing types of perturbation in the problem, for example, corresponding to *stability with respect to initial perturbations,* $\widetilde{\psi}(t) := \psi(t) + \delta\psi(t)$ for a suitable class of perturbations $\delta\psi(t)$ and $\widetilde{y}(t) := \widetilde{y}(t; \widetilde{\psi})$, $y(t) := y(t; \psi)$, and $\mathfrak{d}(P, \widetilde{P}) = \|\delta\psi(t)\|_{[t_0-\tau, t_0]}$. For stability *with respect to persistent perturbations,* one addresses (say) the problems

$$(y + \mu\,\delta y)'(t) = F(t, (y + \mu\,\delta y)(t), (y + \mu\,\delta y)(t - \tau), (y + \mu\,\delta y)'(t - \tau)) + \mu\delta F(t)$$
(3.17)

$(t \in [t_0, \infty))$ for a class of suitable, possibly state-dependent, $\delta F(t)$. For $\mu = 0$ we obtain an unperturbed solution; for $\mu = 1$ we have the solution of a perturbed problem; the associated change δy is the subject of interest when defining stability.

Remark 3.2 *It is possible to express the effect of a perturbation in ψ as a perturbation in the inhomogeneous term. This suggests that there may be a link between stability with respect to initial conditions and stability with respect to persistent perturbations, and for our model equation this is so.*

When defining stability with respect to parameters, the problem $\widetilde{P}$ is obtained from the problem P by substituting $\mathbf{p} + \delta\mathbf{p}$ for $\mathbf{p}$. On occasions, it suffices to develop

---

[xiii]The definition of stability with respect to the time lag requires careful attention to detail, since the initial function will be liable to change if the time-lag changes .

an asymptotic analysis and study the dominant effect – ignoring second-order effects – of small perturbations (leading, for example, to the introduction of sensitivity coefficients; see §3.2).

3.8   Illustrative example of stability

Concepts (1)–(3) are examples of a general notion (there are many refinements) of stability of a solution of an evolutionary problem for an ODE or DDE, or for an NDDE. Let us indicate some stability concepts by reference to

$$y(t) \equiv y(\psi, \rho_{0,1,2}, f; t) \tag{3.18a}$$

$$\{y(t) - \rho_2 y(t - \tau)\}' = \rho_0 y(t) + \rho_1 y(t - \tau) + f(t) \quad (t \in [t_0, T)), \tag{3.18b}$$

$$y(t) = \psi(t) \quad (t \in [t_0 - \tau, t_0). \tag{3.18c}$$

If ψ is perturbed to $\widetilde{\psi}(t) := \psi + \delta\psi$ (τ remaining fixed, and with $\delta\psi$ in some suitable smoothness class) then the size of this perturbation can be measured by some norm $\|\delta\psi\|_{[t_0-\tau,t_0]}$, and stability can be discussed in terms of the boundedness of $\delta y(t)$ for $t \in [t_0, T]$ where $\delta y(t) := \widetilde{y}(t) - y(\psi, \rho_{0,1,2}, f; t)$, with $\widetilde{y}(t) = y(\psi + \delta\psi, \rho_{0,1,2}, f; t)$. If ψ is unchanged but $f(t)$ is perturbed to $\widetilde{f}(t) := f(t) + \delta f(t)$ then we can measure the size of the perturbation by some norm $\|\delta f\|_{[t_0,\infty)}$ and instead define $\delta y(t) := y(\psi, \rho_{0,1,2}, f + \delta f; t) - y(\psi, \rho_{0,1,2}, f; t)$. For a change $\boldsymbol{\delta p}$ in $\mathbf{p} = [\rho_0, \rho_1, \rho_2]$ the size is measured by any vector-norm $\|\boldsymbol{\delta p}\|$. The effect of perturbations can, in principle, be investigated through (3.3) and the corresponding equation in which (with obvious perturbations) $\widetilde{\mathcal{C}}$ replaces $\mathcal{C}$, $\widetilde{\mathcal{K}}$ replaces $\mathcal{K}$, $\widetilde{\mathcal{Y}}$ replaces $\mathcal{Y}$, $\widetilde{\mathfrak{U}}$ replaces $\mathfrak{U}$, $\widetilde{\psi}$ replaces ψ, *etc.*, and an adjustment is made for non-zero t_0. The latter approach gives valuable insight, though alternative methods of analysis of stability may prove to be simpler; the following theorem provides an example.

THEOREM 3.4 *For the problem (3.18), if $\psi, \psi + \delta\psi \in C[t_0 - \tau, t_0]$ then $y(\psi + \delta\psi, \rho_{0,1,2}, f; t) - y(\psi, \rho_{0,1,2}, f; t)$ is bounded and tends to zero provided all the zeros of*

$$\chi(\zeta) := \zeta - \rho_0 - \rho_1 \exp(-\tau\zeta) - \rho_2 \zeta \exp(-\tau\zeta) \tag{3.19}$$

satisfy $$\Re(\zeta) \leq \widehat{\zeta} \text{ where } \widehat{\zeta} < 0. \tag{3.20}$$

Suppose that it is known that $\delta f \in L_1[t_0, \infty)$; then the change δf induces a uniformly bounded change $\delta y(t)$ under the same condition (3.20) on the zeros of (3.19).

The corresponding result for the function $u(t) = y(\tau t)$ (see (1.5)) can be stated by reference to the function $\chi_\natural(\zeta) := \zeta - \tau\rho_0 - \tau\rho_1 \exp(-\zeta) - \rho_2 \zeta \exp(-\zeta)$ replacing χ.

EXAMPLE 3.1 *All solutions of (3.18) are unstable if $|\rho_2| > 1$. A region of asymptotic stability for (3.18) is a set of parameter values for which the zeros of χ satisfy the condition (3.20). Figure 4 displays stability regions for solutions of (3.18) in which τ has been normalized to unity (see Remark 1.1).*
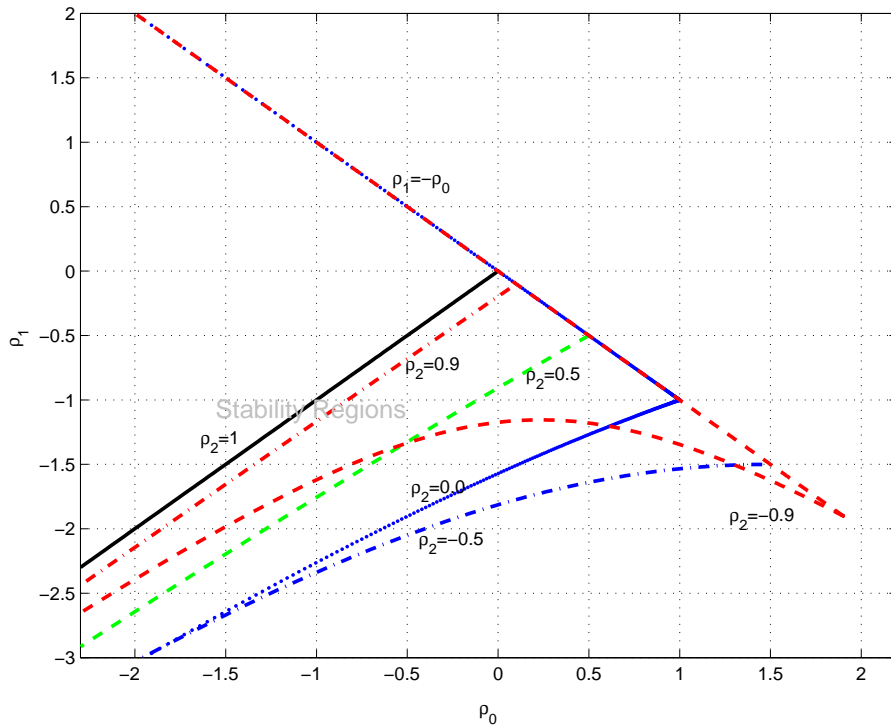
Fig. 4. This figure shows, in the (ρ_0, ρ_1)-plane, the asymptotic stability regions (and their boundaries) for solutions of (3.18) with $\rho_2 = -0.9, -0.5, 0, 0.5, 0.9$ ($\rho_0, \rho_1 \in \mathbb{R}$ and $\tau = 1$); *cf.* [33, Fig.3.1].

3.9 Sensitivity revisited

First-order sensitivity functions and stability analysis each contribute, in rather different ways, a limited understanding of the interaction between parameterized models and solutions, and the recovery of parameter estimates from data.

Let us suppose that a solution is asymptotically stable with respect to a parameter p_ℓ (or, if one prefers, an initial function ψ; in that case the following text must be modified). Due to the asymptotic stability, perturbations in the solution arising from changes in p_ℓ tend to zero as $t \to \infty$. It follows, conversely, that data corresponding to very large values of t can correspond to widely disparate values of p_ℓ, and the estimation of this parameter is therefore difficult if one only has data at large t. On the other hand, we can generalize such conclusions to other values of t if the first-order sensitivity function is small in a neighbourhood of $t = t^*$, the data corresponds to value of t close to t^*, and the perturbations admitted are small. (Of course, one may estimate p_ℓ poorly if one is only interested in re-creating observables near the same value t^*.)

4 Generalizations and future work

Modelling with NDDE models represents an active area of research. The corresponding mathematical techniques require further attention from the applied mathematics community to became a versatile modelling tool capable of adding value to the biology.

Our treatment is not intended to be exhaustive, and we comment on further directions for research. There is a link between stability and the general idea of sensitivity of a solution to perturbations in the problem. A notable omission here is that of an in-depth discussion of adjoint theory for NDDEs.

In the case of sensitivity, one envisages a family of problems, that we denote $\{P(\varepsilon) \,|\, 0 \leq \varepsilon \leq 1\}\}$ where $P(\varepsilon)$ can be considered to be the problem $P + \varepsilon \delta P$, with $P(1) = \widetilde{P}$, say, and (by way of a normalization) $\mathfrak{d}(P, \widetilde{P}) = 1$. If the solution of the problem $P(\varepsilon)$ is denoted $y(\varepsilon; t)$, the first order [xiv] sensitivity is $\lim_{\varepsilon \to 0+} s(\varepsilon; t)$ (provided that the limit exists) where $s(\varepsilon; t) := \{y(\varepsilon; t) - y(0; t)\}/\varepsilon$. It is frequently possible to define the sensitivity function in terms of a Gâteau (or Gateau) derivative. Given smoothness, $\widetilde{y}(t) = y(t) + \int_0^1 s(\varepsilon; t)\,\mathrm{d}\varepsilon$, which may permit insight into stability to be obtained from an investigation of sensitivity.

An interesting question is sensitivity of a solution to perturbations in the initial function (other model features remaining unchanged); for further reading see [13,14,26,40,41]. In this context, the results of the present paper suggest that it could be interesting to revisit those in [40,41] obtained, independently of the work here, by Fathalla Rihan.

We have concentrated here on a hierarchy of *linear* model equations. In general, one would like to consider nonlinear models such as e.g., the generalization of the logistic equation

$$y'(t) = \varrho_0 y(t)\{1 - [\varrho_1 y(t) + \varrho_2 y(t - \tau) + \varrho_3 y'(t - \tau)]\} \tag{4.1}$$

and systems of similar equations. One would also wish to depart from equations that depend upon a single time-lag.

We may generalize the class of equations in our earlier discussion. Given a suitable function u and a value $\tau > 0$, we denote by u_τ the function with values $u_\tau(t) = u(t - \tau)$. Now suppose that $\mathbf{p}$ is a vector-valued parameter. Then a general class of equations comprises those of the form

$$\frac{d}{dt}\left\{y(t) - g_*\Big(\mathbf{p}; t, y(t), y_\tau(t))\Big)\right\} = g^*\Big(\mathbf{p}; t, y(t), y_\tau(t))\Big). \tag{4.2}$$

[xiv]Similar, higher-order sensitivity coefficients can be defined given sufficient smoothness.

To determine a solution of such an NDDE we require an initial condition of the form $y_\tau(0) = \psi(\cdot)$. Generalizing (1.3b), the corresponding *explicit* equations commonly have the form

$$y'(t) = g_\natural\Big(\mathbf{p}; t, y(t), y_\tau(t)), y'_\tau(t))\Big) \ (t \in [0, T]). \tag{4.3}$$

Acknowledgements

References

[1] S.M. Andrew, C.T.H. Baker, G.A. Bocharov, Rival approaches to mathematical modelling in immunology, *J. Comput. Appl. Math.* 205 (2007) 669-686; published online 25 July 2006.

[2] C.T.H. Baker, Retarded differential equations, *J. Comput. Appl. Math.*, 125 (2000) 309-335.

[3] C.T.H. Baker, G.A. Bocharov, Computational aspects of time-lag models of Marchuk type that arise in immunology, *Russian J. Numer. Anal. Math. Modelling*, 20 (2005) 247–262.

[4] C.T.H. Baker, G.A. Bocharov, J.M. Ford, P.M. Lumb, S.J. Norton, C.A.H. Paul, T. Junt, P. Krebs, B. Ludewig, Computational approaches to parameter estimation and model selection in immunology, *J. Comput. Appl. Math.* 184 (2005) 50-76.

[5] C.T.H. Baker, G.A. Bocharov, C.A.H. Paul, F.A. Rihan, Modelling and analysis of time-lags in some basic patterns of cell proliferation. *J. Math. Biol.* 37, No.4, 341-371 (1998).

[6] C.T.H. Baker, G.A. Bocharov, F.A. Rihan, A report on the use of delay differential equations in numerical modelling in the biosciences, *MCCM Tech. Rep.* **343**, ISSN 1360-1725, Univerity of Manchester (1999) 46pp.

[7] C.T.H. Baker, G.A. Bocharov, C.A.H. Paul, Mathematical modelling of the Interleukin-2 T-cell system: A comparative study of approaches based on ordinary and delay differential equations. *Journal of Theoretical Medicine* **2** (1997) 117-128.

[8] C.T.H. Baker, G.A. Bocharov, C.A.H. Paul, F.A. Rihan, Computational modelling with functional differential equations: identification, selection, and sensitivity, *Appl. Numer. Math.*, 53 (2005) 107–129.

[9] C.T.H. Baker, G.A. Bocharov, C.A.H. Paul, F.A. Rihan, Models with delay for cell population dynamics: Identification, selection and analysis. *Appl. Num. Math.* 53 (2005) 107–129.

[10] C.T.H. Baker, P.M. Lumb, On integral equation formulations of a class of evolutionary equations with time-lag, *J. Integral Equations Appl.* 18 (2006) 227–247.

[11] C.T.H. Baker, C.A.H. Paul, Pitfalls in parameter estimation for delay-differential equations, *SIAM J. Sci. Comput.* 18 (1997) 305–314.

[12] C.T.H. Baker, C.A.H. Paul, Discontinuous solutions of neutral delay differential equations. *Appl. Numer. Math.*, 56 (2006) 284–304.

[13] C.T.H. Baker, E.I. Parmuzin, Analysis via integral equations of an identification problem for delay-differential equations, *J. Integral Equations Appl.*, 16 (2004) 111-135 .

[14] C.T.H. Baker, E.I. Parmuzin, Initial function estimation for scalar neutral delay differential equations, submitted for publication.

[15] J.E. Bailey, D.F. Ollis *Biochemical Engineering Fundamentals*, McGraw-Hill Book Company, New York, 1986.

[16] H.T. Banks, B.G. Fitzpatrick, Statistical methods for model comparison in parameter estimation problems for distributed systems. *J. Math. Biol.* **28**, 501-527 (1990).

[17] Y. Bard, *Nonlinear Parameter Estimation* Academic Press, New York, 1974.

[18] D.M. Bates, D.G. Watts, *Nonlinear Regression Analysis and its Applications*, New York: John Wiley & Sons, 1988.

[19] L. Belkoura, M. Dambrine, Y. Orlov, J-P. Richard, Identifiability and identification of linear systems with delays. In *Advances in time-delay systems*, 123–135, Lect. Notes Comput. Sci. Eng., 38, Springer, Berlin, 2004.

[20] A. Bellen, M. Zennaro, *Numerical methods for delay differential equations*, Oxford University Press, New York, 2003.

[21] R. Bellman, K.L. Cooke, *Differential-difference equations*, Academic Press, New York, 1963.

[22] Bocharov G.A., Hadeler K.P., Structured population models, conservation laws, and delay equations, *J. Differ. Equ.*, **168** (2000) 212-237.

[23] G.A. Bocharov, F.A. Rihan, Numerical modelling in biosciences using delay differential equations, *J. Comput. Appl. Math.*, **125** (2000) 183-199.

[24] K.P. Burnham, D.R. Anderson, *Model Selection and Multimodel Inference - a practical information-theoretic approach, 2nd ed.*, Springer-Verlag, New York (2002; Corr. 3rd printing, 2004).

[25] J. Chambers, M. Hill, Fitting nonlinear models: numerical techniques, *Biometrika* **60** (1973) 1-13.

[26] P.C. Das, S.S. Prabhu, Determination of the optimal initial function and the parameters in a system described by delay-differential equations. Journal of Optimization Theory and Applications, 40, (1983) 583-593.

[27] R.D. Driver, Existence and continuous dependence of solutions of a neutral functional-differential equation. Arch. Rational Mech. Anal. 19 (1965) 149–166.

[28] A. Friedman *Foundations of Modern Analysis*, Dover. New York (1982).

[29] P.D. Grünwald, In Jae Myung, M.A. Pitt (Eds.) *Advances in Minimum Description Length Theory and Applications* The MIT Press, Cambridge, MA (2005) 454 pp., 73 illus. [see also `http://homepages.cwi.nl/~pdg/ftp/mdlintro.pdf`]

[30] N. Guglielmi, E. Hairer; RADAR5: a Fortran90 code for the numerical integration of stiff and implicit systems of delay differential equations:
`http//www.unige.ch/math/folks/hairer`.

[31] J.K. Hale, *Theory of functional differential equations*, First edition, 1971; second edition, Springer, New York, 1977.

[32] J.K. Hale, M.A. Cruz, Existence, uniqueness, and continuous dependence for hereditary systems, *Ann. Mat. Pura Appl.*, 85 (1970) 63–81.

[33] V.B. Kolmanovskii and V.R. Nosov, *Stability of Functional Differential Equations*, Academic Press, New York, 1986.

[34] V.B. Kolmanovskiĭ, A.D. Myshkis, *Applied Theory of Functional-differential Equations*, Kluwer Acad. Publ., Dordrecht, 1992.

[35] V.B. Kolmanovskiĭ, A.D. Myshkis, *Introduction to the Theory and Applications of Functional-differential Equations*, Kluwer Acad. Publ., Dordrecht, 1999.

[36] E.M. Landaw, J.J. DiStefano, Multiexponential, multicompartmental, and noncompartmental modeling. II Data analysis and statistical considerations, *Am. J. Physiol.* **5** 665–677.

[37] A.D. Lanterman, Schwarz, Wallace, and Rissanen: Intertwining themes in theories of model selection, *International Statistical Review*, 69 (2001) 185–212.

[38] N. MacDonald, *Time lags in biological models*, Springer, Berlin-New York, 1978.

[39] C.A.H. Paul, A user-guide to `Archi` - An explicit Runge-Kutta code for solving delay and neutral differential equations and parameter estimation problems, *MCCM TR* **283**, *University of Manchester* (1997).

[40] F.A. Rihan, Sensitivity analysis of cell growth dynamics with time lags, *J. Egyptian Math. Soci.* **14** (2006) 91-107.

[41] F.A. Rihan, Sensitivity analysis for dynamic systems with time-lags, *J. Comput. Appl. Math.*, 151 (2003) 445–462.

[42] L.F. Shampine, S. Thompson, Solving DDEs in Matlab, *Appl. Numer. Math.*, 37 (2001) 441458.

[43] L.B. Sheiner and S.L. Beal, Pharmcokinetic parameter estimates from several least squares procedures: superiority of extended least squares, *J. Pharmacokinet Biompharm*, 13 (1985) 185–201.

[44] L. Ljung, *System Identification: Theory for the User*. Prentice-Hall, Inc., Upper Saddle River, N.J., 1987.

[45] J.M. Mitchison, W.S. Vincent, A method of making synchronous cell cultures by density gradient centrifugation. In: I.L. Cameron *et al.*: *Cell Synchrony. Studies in Biosynthetic Regulation,* Academic Press, New York, 1966.

[46] H. L. Van Trees, K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, Wiley-IEEE Press [Wiley Interscience, USA], 2007.

[47] M. Zennaro, The numerical solution of delay differential equations, Summer School on Delay Differential Equations and Applications Dobbiaco (BZ), Italy, June 26th-30th, 2006 `http://www.dmi.units.it/~torelli/dobbiaco06_Zennaro.pdf`