MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume 10, Number 5&6, October & December 2013
doi:10.3934/mbe.2013.10.1301
pp. 1301–1333

THEORETICAL FOUNDATIONS FOR TRADITIONAL AND GENERALIZED SENSITIVITY FUNCTIONS FOR NONLINEAR DELAY DIFFERENTIAL EQUATIONS

H. Thomas Banks, Danielle Robbins and Karyn L. Sutton

Center for Research in Scientific Computation
Center for Quantitative Sciences in Biomedicine
Raleigh, NC 27695-8212, USA

On the Occasion of the 60th Birthday of Carlos Castillo-Chavez

Abstract. In this paper we present new results for differentiability of delay systems with respect to initial conditions and delays. After motivating our results with a wide range of delay examples arising in biology applications, we further note the need for sensitivity functions (both traditional and generalized sensitivity functions), especially in control and estimation problems. We summarize general existence and uniqueness results before turning to our main results on differentiation with respect to delays, etc. Finally we discuss use of our results in the context of estimation problems.

1. Introduction.

There are a wide number of applications where delay equations arise naturally. Moreover, in these applications parameter estimation or inverse problems are ubiquitous. In such problems sensitivity analysis is important in the context of inverse problems, not only in design of experiments, but also in statistical analysis for uncertainty quantification (standard errors, confidence intervals) for parameter estimates. However the theory for sensitivity is not complete and some fundamental issues are yet to be resolved. Here new theoretical fundamental results for sensitivity of delay systems are given; in this context traditional as well as generalized sensitivity functions are formulated. Finally, a discussion of their use in inverse problems along with several computational examples are presented to illustrate their use.

1.1. Delay systems in the biological sciences.

For many years delay equations have been used in biological applications [1, 5, 10, 22, 23, 29, 30, 31, 32, 35, 36, 39, 40, 42, 45, 46, 49, 53, 55, 56, 63]. Very early interest focused around the studies of mechanical systems by Minorsky [58, 59, 60] and slightly later those of Hutchinson [50, 51] in biology. These authors argued that time delays in dynamical systems can produce oscillatory phenomena in an otherwise non-oscillatory system.

In 1948 Hutchinson [50] developed a delay differential equation model, now known as *Hutchinson's equation*, which is the *delayed logistic equation*, to describe the dynamics of a circular causal system. A causal system is a system in which the output depends only on the current and/or past input but not future input. A

2010 *Mathematics Subject Classification.* 34K29, 34K60, 49K40.

Key words and phrases. Delay equations, differentiability with respect to delays, sensitivity, generalized sensitivity, Fisher information matrix.

circular causal system (sometimes encountered in the context of temporal multi-scale models) is any causal system where changes to one part of the system affect another part of the system at a different rate in a manner such that the system does not go extinct. An example of an ecological circular causal system is a parasite-host interaction where a parasite completes its life cycle without killing the host or drastically altering the growth of the host population. The host population can then continue to exist [50, 55]. The delay in this model can represent various naturally occurring phenomena such as the gestation period in a growing population, the life cycle of a parasite, cell cycle delays, etc. Hutchinson's equation (to be used in the numerical illustrations below), its variations and other delay systems have also been used to model physiological control systems as well as numerous other biological processes.

After the early work of Hutchinson there was much subsequent interest in models with delays producing unstable equilibria and possible oscillatory (periodic) behavior. Growth models in which delays played a significant role include those discussed in [32]. Other investigations involved the modeling of gene regulation based on the work of Goodwin [43, 44]. The so-called "Goodwin models" and their numerous variations concerned biochemical pathways in cells, in particular synthesis of proteins controlled via negative feedback or repression. Among the many contributions in this area we mention [22, 23, 56, 57] and the references therein where the focus is on cyclic gene models and the possible existence (or not) of oscillations due to delays attributed to transcription and translation (we remark that similar questions arise more recently in cellular level models of HIV infection pathways [10, 11]). Hadeler [46] in his discussion of several such examples focuses mostly on the destabilization of an equilibrium point and existence of oscillatory solutions in these examples. He discusses the Hutchinson equation and its ability to produce oscillations but also presents the Nicholson blowfly example as an example leading to conclusions quite different from the Hutchinson equation. (The delay here is introduced by the time needed for an egg to become an adult fly, seemingly similar reasoning as that with the Hutchinson equation.) The blowfly equation is formulated with population density as exponential growth and death, with delay in the density dependent birth rate and no delay in the density dependent death rate. In this model, destabilization occurs if the birth term at the equilibrium population density $\bar{N}$ is sufficiently steep or the increase of the death term is sufficiently flat. This is in contrast to the Hutchinson model in which destabilization occurs for sufficiently large delays for any choice of the other parameters. Hadeler also cites a nonlinear delayed restoring force harmonic oscillator model of the sunflower [33, 65] and its geotropic nutations, pointing out that this and many other plants perform complicated movements, and "any deviation from the vertical position leads to compensatory unequal growth of the sides of the stem. The compensatory growth does not stop when the vertical position is achieved, and rather regular oscillations may occur." In a relatively recent book [55], Kuang discusses numerous applications in biology involving nonlinear models with discrete delays including: the effect of incubation delays in a model of malaria epidemic dynamics as presented in [64], the modeling of so-called "dynamic diseases", introduced by Glass and Mackey in 1979 [42], who consider acute physiological disease which appear as an alteration in a control system that is normally periodic, or the onset of an oscillation in an otherwise non-oscillatory system. Kuang also discusses the more recent work of Schuster and Schuster [63] who proposed a model (the logistic equation with delays in both the growth and death rates) to describe Ehrlich Ascites tumor growth in mice. In this regard we mention the papers of Forys and Marciniak-Czocha [39, 40] who summarized efforts in using logistic-type models with delays to describe tumor growth.

Several recent contributions in biological modeling underline the need for sensitivity functions for systems with delays. In [29] Burns, Cliff, and Doughty explain the use of continuous sensitivity equations for a model of the cellular dynamics for Chlamydia Trachomatis while Kappel [53] discusses generalized sensitivities in dynamics of threshold-driven infections. Banks, Banks, and Joyner [9] present a mathematical and statistical framework involving sensitivity for delay systems arising in models for the delayed action of sublethal insecticides in a recent ecological application. None of these presentations give rigorous proofs on the existence of the various Fréchet derivatives [8] that define the sensitivity equations, but the authors of [9] cite some initial theoretical foundations presented in [18, 19], which rely on an abstract theoretical framework presented in [24]. Finally, the efforts in [25] on the dynamics of behavior change in problem drinkers is a modern application in psychology that motivates the need of sensitivity functions for discrete as well as cumulative delayed effects as represented in longitudinal data for patients undergoing therapy.

1.2. **Sensitivity and delay systems.** In both theoretical and qualitative contributions [27, 34, 35, 47], as well as more computational treatments (see for example [5, 6, 7, 8, 12, 13, 21, 52, 54] as well as a large number references therein), a wide range of contributions were made to the literature on delay equations beginning in the 1970s. In some of these early efforts, parameter estimation and control system questions led to the investigation of *traditional sensitivity functions (TSFs)* for delay systems. These TSFs and the recent more general concept of *generalized sensitivity functions (GSFs)* are the focus of our investigations here. In an early paper [14], the authors observed difficulty when estimating the delay and suggested that this could be due to the fact that solutions of DDEs may not always be differentiable with respect to the delays. This of course makes estimation methods such as least squares and maximum likelihood challenging in the case that derivative-based optimization routines are to be used. These authors also suggested the need for a formal theory regarding the existence of sensitivity functions with respect to the delay and among the earliest rigorous results were those of Gibson and Clark [41] and Brewer [28] in their treatment of linear DDEs. By employing semigroups these authors were able to use general representation results to establish the existence of Fréchet derivatives with respect to the parameters for general delay equation initial value problems. As a result of the existence of these Fréchet derivatives, they are able to carefully and rigorously define sensitivity equations with respect to the parameters *including the delay* for general linear systems.

In a more recent report [2], Baker and Rihan *formally* derive sensitivity equations for delay differential equation models, as well as the equations for the sensitivity of parameter estimates with respect to observations (these latter sensitivities are what we shall discuss below as *Generalized Sensitivity Functions (GSFs)*). Baker and Rihan also offer an outline on how to numerically compute both TSFs and GSFs for retarded delay differential equations. While their focus is on computational methods, they also list issues that arise when carrying out parameter estimation in DDEs. These include difficulty in establishing existence of the derivatives of the solution with respect to the parameters and the delays, as well as difficulty in establishing well-posedness for the derived sensitivity equations. Some of these issues are dealt with in a rigorous manner below.

Banks and Bortz [10] were among the first to consider sensitivity with respect to distributional delays. They used sensitivity analysis to show how changes in distributed parameters will affect the solutions of their nonlinear delay differential equation model for HIV progression at the cellular level where intracellular processing delays are distributed across cell populations. The models are validated with what is called *aggregate data* [8].

When deriving the sensitivity equations Banks and Bortz obtain a system of DDEs, which are assumed to be well-posed. In their discussion of well-posedness for these sensitivity equations they assume the delay distributions are differentiable and parameterizable by a mean and standard deviation. In [10] they use theoretical steps (i.e., successive approximations, fixed point theory, Lipschitz continuity, etc.) employed in [7] to prove existence and uniqueness of the resulting sensitivities and sensitivity equations. Motivated by the efforts in [10], Banks and Nguyen [24] develop a rigorous theoretical framework for sensitivity functions for general nonlinear dynamical systems in a Banach space X where the parameters μ are themselves members of another Banach space $\mathcal{M}$. In this setting they consider the sensitivity of solutions x with respect to parameters μ in the following type of abstract nonlinear ordinary differential equations

$$\begin{aligned} \dot{x}(t) &= f(t, x(t), \mu), \ t \geq t_0 \\ x(t_0) &= x_0, \end{aligned} \tag{1}$$

where $f : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ and $\mathcal{M}$ and X are complex Banach spaces. They establish well-posedness for (1), and existence of Fréchet derivatives of the solution $x(t)$ with respect to the parameters μ. As a result, there is a unique solution to the corresponding sensitivity equation

$$\begin{aligned} \dot{y}(t) &= f_x(t, x(t, t_0, x_0, \mu), \mu)y(t) + f_\mu(t, x(t, t_0, x_0, \mu), \mu), \quad t \geq t_0 \\ y(t_0) &= 0, \end{aligned} \tag{2}$$

where $y(t) = \dfrac{\partial x(t)}{\partial \mu}$. In [24] Banks and Nguyen provide rigorous theoretical sensitivity results for the DDE example for HIV dynamics with measure dependent or distributional parameters given in [10]; however they only present results for the the sensitivity with respect to absolutely continuous probability distributions for the delay. In subsequent efforts [18, 19] a rigorous theoretical foundation is developed for sensitivity theory using *directional derivatives* where the parameter space $\mathcal{M}$ is taken as the convex metric space of probability measures (including discrete, continuous or convex combinations thereof) taken with the Prohorov metric topology [8]. Below we give new results for sensitivity with respect to discrete delays. The proofs, while quite tedious, continue with an adaption of the well known ideas for existence and uniqueness of the Fréchet derivative with respect to the delay in nonlinear DDE as employed in [10, 18, 19, 24].

After summarizing recent theoretical results on differentiability with respect to parameters, initial conditions and discrete delays, we discuss both traditional and generalized sensitivity functions with respect to the same quantities. Finally, to illustrate computationally the use of these sensitivity functions, we turn to two classical examples: the Hutchinson delayed growth model and the harmonic oscillator with delays introduced many years ago by Minorsky.

2. Solutions and approximations.

We first summarize existence and uniqueness for general nonlinear nonautonomous dynamical systems involving delays of the form

$$\dot{x}(t) = G(t, x(t), x_t, x(t-\tau_1), \ldots, x(t-\tau_m), \theta) + G_2(t), \quad 0 \le t \le T, \tag{3}$$
$$x(t) = \phi(t) \quad -r \le t \le 0,$$

where $G = G(t, \eta, \psi, y_1, \ldots, y_m) : [0,T] \times X \times \mathbb{R}^{nm} \times \mathbb{R}^p \to \mathbb{R}^n$. Here $X = \mathbb{R}^n \times L_2(-r, 0; \mathbb{R}^n)$, $0 < \tau_1 < \ldots < \tau_m = r$, x_t denotes the usual function $x_t(\xi) = x(t+\xi)$, $-r \le \xi \le 0$, while $x(t)$ is a point in $\mathbb{R}^n$, and $\phi \in H^1(-r, 0)$. The function G_2 is a time dependent perturbation (e.g., a control input).

We turn to the mathematical aspects of these nonlinear FDE systems and present an outline of the necessary mathematical foundations. First we describe the conversion of the nonlinear FDE system to an abstract evolution equation (AEE) as well as provide existence and uniqueness results for a solution to the FDE. One can use the ideas of a linear semigroup framework, in which approximation of linear delay systems has been developed, as a basis for a wide class of nonlinear delay system approximations. Details in this direction can be found in the early work [6, 7, 52] which is a direct extension of the results in [12, 13, 21] to nonlinear delay systems.

We shall make use of the following hypotheses throughout our presentation.

(H1) The function G satisfies a global Lipschitz condition:

$$|G(t, \eta, \psi, y_1, \ldots, y_m) - G(t, \tilde{\eta}, \tilde{\psi}, w_1, \ldots, w_m)| \le K \left(|\eta - \tilde{\eta}| + |\psi - \tilde{\psi}| + \sum_{i=1}^{m} |y_i - w_i| \right)$$

for some fixed constant K and all $(\eta, \psi, y_1, \ldots, y_m)$, $(\tilde{\eta}, \tilde{\psi}, w_1, \ldots, w_m)$ in $X \times \mathbb{R}^{nm}$ uniformly in t.

(H2) The function $G : [0,T] \times X \times \mathbb{R}^{nm} \to \mathbb{R}^n$ is differentiable.

Remark 1. If we define the function $g : [0,T] \times \mathbb{R}^n \times C(-r, 0; \mathbb{R}^n) \subset X \to \mathbb{R}^n$ given by

$$g(t, x) = G(t, \eta, \psi) = G(t, \eta, \psi, \psi(-\tau_1), \ldots, \psi(-\tau_m)), \tag{4}$$

we observe that even though G satisfies (H1), g will not satisfy a continuity hypothesis on its domain in the X norm.

Letting $z(t) = (x(t), x_t) \in X$, where the Hilbert space X has the inner product

$$\langle (\eta, \phi), (\zeta, \psi) \rangle_X = < \eta, \zeta >_{\mathbb{R}^n} + \int_{-r}^{0} \phi(\xi) \psi(\xi) d\xi, \tag{5}$$

we define the nonlinear operator $\mathcal{A}(t) : \mathcal{D}(\mathcal{A}) \subset X \to X$ by

$$\mathcal{D}(\mathcal{A}) \equiv \{(\psi(0), \psi) \mid \psi \in H^1(-r, 0)\}$$

$$\mathcal{A}(t)(\psi(0), \psi) = (g(t, \psi(0), \psi), D\psi)$$

where here $D\psi = \psi'$. Then the functional differential equation (FDE) (3) can be formulated as

$$\begin{aligned} \dot{z}(t) &= \mathcal{A}(t) z(t) + G_2(t) \\ z(0) &= z_0, \end{aligned} \tag{6}$$

where $z_0 = (\phi(0), \phi)$ is the initial condition.

Theorem 2. *Assume that (H1) holds and let* $z(t;\phi,G_2) = (x(t;\phi,G_2), x_t(\phi,G_2))$*, where* x *is the solution of* (3) *corresponding to* $\phi \in H^1$*,* $G_2 \in L_2$*. Then for* $\zeta = (\phi(0),\phi)$*,* $z(t;\phi,G_2)$ *is the unique solution on* $[0,T]$ *of*

$$z(t) = \zeta + \int_0^t [\mathcal{A}(\sigma)z(\sigma) + (G_2(\sigma),0)]d\sigma. \tag{7}$$

Furthermore, $G_2 \to z(t;\phi,G_2)$ *is weakly sequentially continuous from* L_2 *(with weak topology) to* X *(with strong topology).*

These results can be established in one of several ways [7] including fixed point theorem arguments or Picard iteration arguments; either approach can be used to establish existence, uniqueness and continuous dependence of the solution of (7). More complete discussions can be found in [8] and the detailed references given therein. We remark that our condition (H1) is a global version of the hypothesis of Kappel and Schappacher in [54], so that in the autonomous case their results also yield immediately the desired existence and uniqueness result for (3). We note that Theorem 2 can be readily extended to the case $z_0 = (x_0,\phi)$ where $x_0 \neq phi(0)$, necessarily.

3. Continuous dependence and differentiability.

To establish continuous dependence in parameters and differentiability with respect to model parameters, initial conditions, and a time delay (the latter not previously done elsewhere to the authors' knowledge), we focus for ease of arguments on a simple case with a nonlinear autonomous system with only one discrete delay (the multiple delay case is handled in a completely similar manner) of the form

$$\frac{dx(t)}{dt} = G(x(t), x(t-\tau), \theta), \quad t \in [0,T], \tag{8}$$

$$x(\xi) = \begin{cases} \phi(\xi), & -\tau \le \xi < 0, \\ x_0, & \xi = 0, \end{cases} \tag{9}$$

where $z = (x_0,\phi) \in Z \equiv \mathbb{R}^n \times L^2(-\tau,0;\mathbb{R}^n)$, $x(t) = x(t;z,\tau,\theta) \in \mathbb{R}^n$, $\tau > 0$ and $\theta \in \mathbb{R}^p$. While we consider here the case of finite dimensional model parameters, similar sensitivity results also hold in a more general case when parameters and delays are distributed, and hence infinite dimensional. Some of these results for infinite dimensional systems are presented in [18, 19]. Once established, these results allow us to study traditional and generalized sensitivity functions, where sensitivity is considered with respect to these three quantities (delays, initial conditions, and parameters). In practice these quantities are often unknown and may need to be estimated from observed or experimental data in inverse problem formulations, e.g., see [8]. The use of sensitivity functions can aid in such endeavors. We begin by considering continuous dependence of solutions $x(t)$ on model parameters θ. We remark that some older delay equation results on continuous dependence and differentiability (with respect to parameters) can be found in [47, 48] although the specific conditions and results with respect to differentiability with respect to discrete delays and the corresponding rigorous development of sensitivity equations are in fact (to our knowledge) new.

Lemma 3. *Let* $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ *and for* $\theta = \theta_0$*, let* $x(t;z,\tau,\theta_0)$ *be a solution of* (8) - (9) *for* $t \in [0,T]$*. Assume that*

$$\lim_{\theta\to\theta_0} G(x,\tilde{x},\theta) = G(x,\tilde{x},\theta_0), \tag{10}$$

uniformly in x and $\tilde{x}$. For $(x_1, \tilde{x}_1, \theta), (x_2, \tilde{x}_2, \theta) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p$ assume that

$$|G(x_1, \tilde{x_1}, \theta) - G(x_2, \tilde{x_2}, \theta)| \leq C_1|x_1 - x_2| + C_2|\tilde{x_1} - \tilde{x_2}| \tag{11}$$

where $C_j > 0$ are constants for $j = 1, 2$. Then the initial value problem (IVP) (8) - (9) has a unique solution $x(t; z, \tau, \theta)$ that satisfies

$$\lim_{\theta \to \theta_0} x(t; z, \tau, \theta) = x(t; z, \tau, \theta_0), \textit{uniformly in } t \in [0, T].$$

A proof of this standard lemma can be found in [26] and is therefore omitted.

One can use similar arguments to prove similar results for continuity of solutions with respect to initial conditions z and delays τ. We shall not do so, but turn next to the main theoretical results of this paper, the differentiability of solutions with respect to model parameters, initial conditions, and the time delay τ. The proofs of these subsequent theorems will involve the use of a mean value theorem, a version of which we include here stated as a simple lemma.

Lemma 4. *Let $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ and $z = (x_0, \phi) \in Z = \mathbb{R}^n \times L^2(-\tau, 0; \mathbb{R}^n)$.*

(i) If the Fréchet derivatives $G_x(x, \tilde{x}, \theta)$ and $G_{\tilde{x}}(x, \tilde{x}, \theta)$ exist and are continuous for $x, \tilde{x} \in \mathbb{R}^n$, then for $x_1, x_2, \tilde{x}_1, \tilde{x}_2 \in \mathbb{R}^n$ and $\theta \in \mathbb{R}^p$, $t \geq 0$,

$$\begin{aligned} &G(x_1, \tilde{x}_1, \theta) - G(x_2, \tilde{x}_2, \theta) \\ &\quad = \int_0^1 \{G_x(sx_1 + (1-s)x_2, s\tilde{x}_1 + (1-s)\tilde{x}_2, \theta)(x_1 - x_2) \\ &\qquad + G_{\tilde{x}}(sx_1 + (1-s)x_2, s\tilde{x}_1 + (1-s)\tilde{x}_2, \theta)(\tilde{x}_1 - \tilde{x}_2)\} ds. \end{aligned}$$

(ii) If the Fréchet derivatives $G_\theta(x, \tilde{x}, \theta)$ exist and are continuous for $x, \tilde{x} \in \mathbb{R}^n$, then for $x, \tilde{x} \in \mathbb{R}^n$ and $\theta_1, \theta_2 \in \mathbb{R}^p$, $t \geq 0$,

$$G(x, \tilde{x}, \theta_1) - G(x, \tilde{x}, \theta_2) = \int_0^1 G_\theta(x, \tilde{x}, s\theta_1 + (1-s)\theta_2)(\theta_1 - \theta_2) ds.$$

(iii) Suppose $x(t, z, \tau, \theta)$ is a solution of (8) - (9), which is continuous in $\mathbb{R}^n$, and continuous and continuously differentiable for $z \in Z$, such that $D_z x(t; \cdot) \in \mathcal{L}(Z, \mathbb{R}^n)$. Then for $z_1, z_2 \in Z$, and fixed $\tau \in \mathbb{R}$, $\theta \in \mathbb{R}^p$, for $t \in [0, T]$,

$$x(t; z_1, \tau, \theta) - x(t; z_2, \tau, \theta) = \int_0^1 D_z x(t; sz_1 + (1-s)z_2)[z_1 - z_2] ds.$$

Proof. We consider *(i)*. Let

$$H_1(s) = G(sx_1 + (1-s)x_2, s\tilde{x}_1 + (1-s)\tilde{x}_2, \theta), \quad 0 < s \leq 1.$$

Using the chain rule with Fréchet derivatives, we have

$$\begin{aligned} H_1'(s) = &G_x(sx_1 + (1-s)x_2, s\tilde{x}_1 + (1-s)\tilde{x}_2, \theta)(x_1 - x_2) \\ &+ G_{\tilde{x}}(sx_1 + (1-s)x_2, s\tilde{x}_1 + (1-s)\tilde{x}_2, \theta)(\tilde{x}_1 - \tilde{x}_2). \end{aligned}$$

If we integrate ${H'}_1(s)$ for $s \in (0, 1]$, we have $H_1(1) - H_1(0)$ which is $G(x_1, \tilde{x}_1, \theta) - G(x_2, \tilde{x}_2, \theta)$. Thus the statement in *(i)* is shown. The proof of *(ii)* and *(iii)* are similar to the proof of *(i)* , and thus are omitted. □

Next we turn to differentiability of solutions of the general system (8)–(9) with respect to model parameters θ in the following theorem. The proof of Theorem 5 is excluded but can be found in [62]. (Moreover, the proof of Theorem 7 given below can easily be followed to give that of Theorem 5.) Thus without further discussion, we proceed to Theorem 6, in which we establish differentiability of the solutions with respect to the initial conditions $z = (x_0, \phi)$.

Theorem 5. *Suppose that $G(x, \tilde{x}, \theta)$ has continuous Fréchet derivatives $G_x, G_{\tilde{x}}, G_\theta$ such that $|G_x| \le M_0$, $|G_{\tilde{x}}| \le M_1$, and $|G_\theta| \le M_2$ for constants $M_j, j = 0, 1, 2$. Then the Fréchet derivative $y_1(t) = \dfrac{\partial x(t)}{\partial \theta} \in \mathbb{R}^{n\times p}$ exists and is the unique solution of*

$$\begin{aligned} \dot{y}_1(t) &= G_x(x(t), x(t-\tau), \theta)y_1(t) + G_{\tilde{x}}(x(t), x(t-\tau), \theta)y_1(t-\tau) \\ &\quad + G_\theta(x(t), x(t-\tau), \theta), \\ y_1(\xi) &= 0 \quad -\tau \le \xi \le 0. \end{aligned} \tag{12}$$

Theorem 6. *Suppose the function $G(x, \tilde{x}, \theta)$ of (8) has continuous Fréchet derivatives $G_x(x, \tilde{x}, \theta)$, $G_{\tilde{x}}(x, \tilde{x}, \theta)$, with respect to x and $\tilde{x}$, with $|G_x| \le M_0$, $|G_{\tilde{x}}| \le M_1$, for some constants $M_j > 0$ for $j = 0, 1$. Then the Fréchet derivative*

$$y_2(t) = \frac{\partial}{\partial z} x(t; z, \theta)$$

exists with $y_2(t)[\cdot] \in \mathcal{L}(Z, \mathbb{R}^n)$ (recall $z = (x_0, \phi) \in Z = \mathbb{R}^n \times L^2(-\tau, 0; \mathbb{R}^n)$), and satisfies the equation for $h = (h_0, \tilde{h}) \in Z$ the linear delay differential equation

$$\dot{y}_2(t)[h] = G_x(x(t), x(t-\tau), \theta)y_2(t)[h] + G_{\tilde{x}}(x(t), x(t-\tau), \theta)y_2(t-\tau)[h], \quad t > 0, \tag{13}$$

$$y_2(\xi)[h] = \begin{cases} \tilde{h}(\xi), & -\tau \le \xi < 0, \\ h_0, & \xi = 0. \end{cases}$$

Proof. For a given solution x of (8)–(9) we recognize that (13) is a nonautonomous linear delay differential equation with continuous time dependent coefficients for which an existence and uniqueness theory has been available for many decades [3, 4, 47]. Let $y_2(t; h)$ denote this unique solution of (13). Then by the variation of parameters representation for linear nonautonomous delay systems given in for example [4], we readily argue that $h \to y_2(t; h)$ is a linear functional in h and hence we will denote it by $y_2(t)[h]$ and observe that $y_2(t) = y_2(t)[\,\cdot\,] \in \mathcal{L}(Z, \mathbb{R}^n)$.

For fixed $\tau \in \mathbb{R}$, $\theta \in \mathbb{R}^p$, and $t \in [0, T]$, let $h \in Z$, and $m_1(t, z, h) = x(t; z + h) - x(t; z)$, which can be written as

$$m_1(t, z, h) = \int_0^t \{G(x(s; z+h), x(s-\tau; z+h), \theta) - G(x(s; z), x(s-\tau; z), \theta)\} ds.$$

With Fréchet differentiability of G with respect to $x, \tilde{x} \in \mathbb{R}^n$ and for $z \in Z$, we have

$$\begin{aligned} &m_1(t, z, h) \\ &\quad = \int_0^t \{G_x(x(s; z), x(s-\tau; z), \theta)[x(s; z+h) - x(s; z)] + w_1(s, m_1(s, z, h)) \\ &\qquad + G_{\tilde{x}}(x(s; z), x(s-\tau; z), \theta)[x(s-\tau; z+h) - x(s-\tau; z)] \\ &\qquad + w_1(s-\tau, m_1(s-\tau, z, h)\}\, ds, \end{aligned}$$

where $\dfrac{|w_1(t, m_1(t, z, h))|}{|m_1(t, z, h)|} \to 0$ as $|m_1(t, z, h)|$ approaches zero. We define $b_1(t, h)$ as $\dfrac{|w_1(t, m_1(t, z, h))|}{|m_1(t, z, h)|}$, so $b_1(t, h) \to 0$ uniformly in t as $|h| \to 0$. Then to argue that

$\dfrac{\partial}{\partial z}x(t;z,\theta) \in \mathcal{L}(Z,\mathbb{R}^n)$ exists and is a solution for (13), we only need to show that

$$\frac{|m_1(t,z,h) - y_2(t)[h]|}{|h|} \to 0$$

as $|h| \to 0$.

By definition

$$\begin{aligned}
&\frac{|m_1(t,z,h) - y_2(t)[h]|}{|h|} \\
&\quad = \frac{1}{|h|}\Big| \int_0^t \{G_x(x(s;z), x(s-\tau;z),\theta)[x(s;z+h) - x(s;z)] \\
&\qquad\quad + G_{\tilde{x}}(x(s;z), x(s-\tau;z),\theta)[x(s-\tau;z+h) - x(s-\tau;z)] \\
&\qquad\quad + w_1(s, m_1(s,z,h)) + w_1(s-\tau, m_1(s-\tau,z,h)) \\
&\qquad\quad - (G_x(x(s,z), x(s-\tau,z),\theta)y_2(s)[h] \\
&\qquad\quad + G_{\tilde{x}}(x(s;z), x(s-\tau;z),\theta)y_2(s-\tau)[h])\}ds\Big|.
\end{aligned}$$

We turn to the term $w_1(s-\tau, m_1(s-\tau,z,h))$, and make the change of variables $\xi = s-\tau$, so that

$$\begin{aligned}
&\int_0^t |w_1(s-\tau, m_1(s-\tau,z,h))|ds \\
&\qquad = \int_{-\tau}^0 |w_1(\xi, m_1(\xi,z,h))|d\xi + \int_0^{t-\tau} |w_1(\xi, m_1(\xi,z,h))|d\xi \\
&\qquad \le \int_{-\tau}^0 |w_1(\xi,h)|d\xi + \int_0^t |w_1(\xi, m_1(\xi,z,h))|d\xi,
\end{aligned}$$

since $m_1(\xi,z,h) = h$ for $\xi \in [-\tau,0]$. Using this and the bounds on the Fréchet derivatives $|G_x|$ and $|G_{\tilde{x}}|$ we have

$$\begin{aligned}
&\frac{|m_1(t,z,h) - y_2(t)[h]|}{|h|} \\
&\quad \le \frac{1}{|h|}\Big[\int_0^t \{M_0|m_1(s,z,h) - y_2(s)[h]| + M_1|m_1(s-\tau,z,h) - y_2(s-\tau)[h]| \\
&\qquad + 2|w_1(s, m_1(s,z,h))|\,\}ds\Big] + \int_{-\tau}^0 b_0(\xi,h)d\xi,
\end{aligned}$$

where $b_0(\xi,h) = \dfrac{|w_1(\xi,h)|}{|h|} \to 0$ uniformly in $\xi \in [-\tau,0]$ as $h \to 0$.

Defining the difference $\Delta(t) = |m_1(t,z,h) - y_2(t)[h]|$ and noting that $\Delta(\xi) = 0$ for $-\tau < \xi < 0$, we find

$$\begin{aligned}\frac{\Delta(t)}{|h|} &\\ &\leq \frac{1}{|h|}\left\{\int_0^t \{(M_0+M_1)\Delta(s))\}ds + \int_0^t 2|w_1(s,m_1(s,z,h))|ds\right\} + \int_{-\tau}^0 b_0(\xi,h)d\xi \\ &= \frac{1}{|h|}\left\{\int_0^t \{(M_0+M_1)\Delta(s))\}ds\right\} + \int_{-\tau}^0 b_0(\xi,h)d\xi \\ &\quad + \frac{1}{|h|}\left\{\int_0^t \{2\frac{|w_1(s,m_1(s,z,h))|}{|m_1(s,z,h)|}\Delta(s) + 2\frac{|w_1(s,m_1(s,z,h))|}{|m_1(s,z,h)|}|y_2(s)[h]|\}ds\right\}.\end{aligned} \tag{14}$$

It follows that

$$\frac{\Delta(t)}{|h|} \leq \int_0^t \{(M_0+M_1+2b_1(s,h))\frac{\Delta(s)}{|h|}\}ds + \int_0^t K\, b_1(s,h)ds + \int_{-\tau}^0 b_0(\xi,h)d\xi. \tag{15}$$

Using the fact that $b_0(\xi,h),\ b_1(t,h) \to 0$ as $h \to 0$ and Gronwall's inequality, we have

$$\begin{aligned}\lim_{|h|\to 0} &\frac{|m_1(t,z,h) - y_2(t)[h]|}{|h|} \\ &\leq \lim_{|h|\to 0}\left[\int_0^T K\, b_1(s,h)ds + \int_{-\tau}^0 b_0(\xi,h)d\xi\right]\left(e^{(M_0+M_1)T+\int_0^T b_1(s,h)ds}\right) \\ &= 0,\end{aligned}$$

which completes the proof. □

Theorem 7. *Suppose that $G(x,\tilde{x},\theta)$ has continuous Fréchet derivatives $G_x, G_{\tilde{x}}$ such that $|G_x| \leq M_0$, and $|G_{\tilde{x}}| \leq M_1$ and suppose that the solution x of (8)-(9) satisfies $x \in H^{1,\infty}(-\tau,T;\mathbb{R}^n)$, for $0 < \tau < r$ for fixed $r > 0$. Then the Fréchet derivative $y_3(t) = \dfrac{\partial x(t)}{\partial \tau} \in \mathbb{R}^n$ exists and is the unique solution for*

$$\begin{aligned}&\dot{y}_3(t) = G_x(x(t),x(t-\tau),\theta)y_3(t) + G_{\tilde{x}}(x(t),x(t-\tau),\theta)[y_3(t-\tau) - \dot{x}(t-\tau)] \quad (16)\\ &y_3(\xi) = 0, \quad -\tau \leq \xi \leq 0.\end{aligned}$$

Moreover, $\dfrac{\partial x(t)}{\partial \tau}$ is continuous in θ and, if $x \in C^1(-\tau,T;\mathbb{R}^n)$ it is also continuous in τ.

Proof. We reformulate (16) as a Cauchy problem on the state space $Z = \mathbb{R}^n \times L^2(-r,0;\mathbb{R}^n)$ with the norm $|(\eta,\phi)|^2 = |\eta|^2 + \int_{-r}^0 |\phi(s)|^2 ds$. One may then consider solutions of the system for τ satisfying $-r < -\tau < 0$.

Let $y_3(t,\tau)$ be the solution to (16) (we suppress notation indicating the dependence of solutions on θ). Again existence and uniqueness follow from classical results [3, 4, 47] given that $x \in H^{1,\infty}(-\tau,T;\mathbb{R}^n)$. Then for $t > 0$ we define $y_{3t}(\cdot) \in L^2(-\tau,0;\mathbb{R}^n)$ by the past history $y_{3t}(\xi,\tau) = y_3(t+\xi,\tau)$, $\quad -\tau < \xi < 0$. If

$z_1(t,\tau) = (y_3(t,\tau), y_{3t}(\cdot,\tau))^T$, then it can be shown [8] that $z_1(t,\tau)$ is the solution to the abstract Cauchy problem

$$\frac{dz_1(t)}{dt} = A(t,\tau)z_1(t,\tau) \tag{17}$$
$$z_1(0,\tau) = (0,0)^T \in Z_1,$$

where $\mathcal{D}(A(t,\tau)) = \{(\eta,\phi(\cdot))^T : \eta \in \mathbb{R}^n, \phi(\cdot) \in H^1(-\tau,0;\mathbb{R}^n), \phi(0)=\eta\}$, and

$$A(t,\tau)\begin{bmatrix}\eta \\ \phi(\cdot)\end{bmatrix} = \begin{bmatrix}G_x(x(t),x(t-\tau),\theta)\eta + G_{\tilde{x}}(x(t),x(t-\tau),\theta)[\phi(-\tau) - \dot{x}(t-\tau)] \\ \phi^{'}(\cdot)\end{bmatrix}.$$

Note that $A(t,\tau)$ is a vector affine operator on $z_1(t) = (y_3(t), y_{3_t}(\cdot))^T$. Thus we proceed to argue that $\dfrac{\partial x}{\partial \tau}$ exists and satisfies (16) (or (17)).

To argue that $\dfrac{\partial x}{\partial \tau}$ exists and satisfies either equivalent system (16) or (17), we define $m_2(t,\tau,h) = x(t,\tau+h) - x(t,\tau) = x_h(t) - x(t)$ where $x_h(t)$ denotes a solution to the system

$$\frac{dx_h(t)}{dt} = G(x_h(t), x_h(t-(\tau+h)),\theta) \quad t>0$$
$$x(\xi) = \begin{cases}\phi(\xi) & -(\tau+h) \le \xi < 0 \\ x_{h0} & \xi = 0\end{cases}.$$

Then to argue that $\dfrac{\partial x}{\partial \tau}$ exists and equals y_3, we prove

$$\frac{|m_2(t,\tau,h) - y_3(t)h|}{|h|} \to 0.$$

In the remainder of the proof, the dependence of G on the model parameters θ is suppressed. The difference $m_2(t,\tau,h)$ is

$$\begin{aligned}
m_2(t,\tau,h) &= \int_0^t [G(x_h(s), x_h(s-(\tau+h))) - G(x(s),x(s-\tau))]\,ds \\
&= \int_0^t \{G(x_h(s), x_h(s-(\tau+h))) - G(x(s), x_h(s-(\tau+h))) \\
&\quad + G(x(s), x_h(s-(\tau+h))) - G(x(s),x(s-\tau))\}ds \\
&= \int_0^t \left\{\int_0^1 G_x(rx_h(s) + (1-r)x(s), x_h(s-(\tau+h)))(x_h(s)-x(s))dr \right.\\
&\quad + \int_0^1 G_{\tilde{x}}(x(s), rx_h(s-(\tau+h)) + (1-r)x(s-\tau)) \\
&\qquad (x_h(s-(\tau+h)) - x(s-\tau))dr\Big\}\,ds.
\end{aligned}$$

Then we have

$$\begin{aligned}
&m_2(t,\tau,h) \\
&\quad = \int_0^t \{G_x(x(s),x(s-\tau))m_2(s,\tau,h) + G_{\tilde{x}}(x(s),x(s-\tau))m_2(s-\tau,\tau,h) \\
&\qquad -G_{\tilde{x}}(x(s),x(s-\tau))[x(s-\tau) - x(s-(\tau+h))] + w_2(s, m_2(s,\tau,h)) \\
&\qquad +w_2(s-\tau, m_2(s-\tau,\tau,h)) + w_3(s-\tau,h)\}\,ds,
\end{aligned}$$

where $\dfrac{|w_2(t,m_2(t,\tau,h))|}{|m_2(t,\tau,h)|} \to 0$ as $m_2(t,\tau,h) \to 0$ and $\dfrac{|w_3(t,x,h)|}{|h|} \to 0$ as $[x(t) - x(t-h)] \to 0$, i.e., as $h \to 0$. As in the previous proof we define a ratio

$$b_2(t,h) = \frac{|w_2(t,m_2(t,\tau,h))|}{|m_2(t,\tau,h)|}$$

as well as

$$b_3(t,h) = \frac{|w_3(t,h)|}{|h|},$$

and observe that in fact $b_2(t,h),\ b_3(t,h) \to 0$ uniformly in t as $h \to 0$.

Next we consider the difference $\Delta(t,\tau,h) = |m_2(t,\tau,h) - y_3(t)h|$ and find from above that

$$\begin{aligned}
&\frac{\Delta(t,\tau,h)}{|h|}\\
&= \frac{1}{|h|}\int_0^t \{G_x(x(s),x(s-\tau))m_2(s,\tau,h) + G_{\tilde{x}}(x(s),x(s-\tau))m_2(s-\tau,\tau,h)\\
&\qquad -G_{\tilde{x}}(x(s),x(s-\tau))[x(s-\tau) - x(s-(\tau+h))] + w_2(s,m_2(s,\tau,h))\\
&\qquad +w_2(s-\tau,m_2(s-\tau,\tau,h)) + w_3(s-\tau,h)\\
&\qquad -h\,[G_x(x(s),x(s-\tau))y_3(s) + G_{\tilde{x}}(x(s),x(s-\tau))y_3(s-\tau)\\
&\qquad -G_{\tilde{x}}(x(s),x(s-\tau))\dot{x}(t-\tau)]\}\,ds,\\
&\le \int_0^t \frac{|G_x(x(s),x(s-\tau))|\Delta(s,\tau,h)}{|h|}ds\\
&\quad + \int_0^t \frac{|G_{\tilde{x}}(x(s),x(s-\tau))|\Delta(s-\tau,\tau,h)}{|h|}ds\\
&\quad + \int_0^t \frac{|G_{\tilde{x}}(x(s),x(s-\tau))[x(s-\tau) - x(s-(\tau+h)] - \dot{x}(s-\tau)h|}{|h|}ds\\
&\quad + \int_0^t \left\{\frac{|w_2(s,m_2(s,\tau,h))|}{|h|} + \frac{|w_2(s-\tau,m_2(s-\tau,\tau,h))|}{|h|} + \frac{|w_3(s-\tau,h)|}{|h|}\right\}ds\\
&\le \int_0^t (M_0+M_1)\frac{\Delta(s,\tau,h)}{|h|}ds + \int_0^t M_1 b_4(s,h)ds\\
&\quad + \int_0^t \left\{2\frac{|w_2(s,m_2(s,\tau,h))|}{|m_2(s,\tau,h)|}\frac{\Delta(s,\tau,h)}{|h|} + 2\frac{|w_2(s,m_2(s,\tau,h))|}{|m_2(s,\tau,h)|}|y_3(s)|\right.\\
&\qquad\quad \left. + \frac{|w_3(s,h)|}{|h|}\right\}ds\\
&\le \int_0^t (M_0+M_1+2b_2(s,h))\frac{\Delta(s,\tau,h)}{|h|}ds\\
&\qquad + \int_0^t \{M_1 b_4(s,h) + 2b_2(s,h)|y_3(s)| + b_3(s,h)\}ds,
\end{aligned}$$

where $b_4(s,h) = \dfrac{|\dot{x}(s-\tau)h - [x(s-\tau) - x(s-(\tau+h))]|}{|h|}$. Thus we find

$$\frac{\Delta(t,\tau,h)}{|h|} \le \int_0^t \{M_0 + M_1 + 2b_2(s,h)\} \frac{\Delta(s,\tau,h)}{|h|} ds + \int_0^T \{M_1 b_4(s,h) + b_3(s,h) + 2K_3 b_2(s,h)\} ds.$$

Using $x \in H^1$, we have that $\lim_{h\to 0} \dfrac{[x(t-\tau) - x(t-(\tau+h))]}{h} = \dot{x}(t-\tau)$ in $L_2(0,T)$, so $b_4(s,h) \to 0$ as $|h| \to 0$. Thus we have that $K(h) \to 0$ as $h \to 0$ where

$$K(h) = \int_0^T \{M_1 b_4(s,h) + b_3(s,h) + 2K_3 b_2(s,h)\} ds.$$

Applying Gronwall's inequality again, we have

$$\lim_{|h|\to 0} \frac{|m_2(t,\tau,h) - y_3(t)h|}{|h|} \le \lim_{|h|\to 0} \{K(h)\} e^{\int_0^T \{M_0 + M_1 + 2b_2(s,h)\} ds} = 0.$$

To argue that the solution to (17) depends continuously on θ, let $\Delta_3(t;\tau,\theta,\theta_0) = z_1(t,\tau,\theta) - z_1(t,\tau,\theta_0)$, where now we need to express explicitly the dependence of solutions on θ. This difference satisfies

$$\begin{aligned}
|\Delta_3(t;\tau,\theta,\theta_0)| =& |z_1(t,\tau,\theta) - z_1(t,\tau,\theta_0)|, \\
\le& \int_0^t |A(s,\tau,\theta) z_1(s,\tau,\theta) - A(s,\tau,\theta_0) z_1(s,\tau,\theta_0)| ds, \\
\le& \int_0^t \{|A(s,\tau,\theta) z_1(s,\tau,\theta) - A(s,\tau,\theta) z_1(s,\tau,\theta_0)| \\
&+ |A(s,\tau,\theta) z_1(s,\tau,\theta_0) - A(s,\tau,\theta_0) z_1(s,\tau,\theta_0)|\} ds, \\
\le& \int_0^t |A(s,\tau,\theta) \Delta_3(s;\tau,\theta,\theta_0)| ds \\
&+ \int_0^t |(A(s,\tau,\theta) - A(s,\tau,\theta_0)) z_1(s,\tau,\theta_0)| ds, \\
\le& Q \int_0^t |\Delta_3(s;\tau,\theta,\theta_0)| ds + r(T;\tau,\theta),
\end{aligned} \tag{18}$$

where $r(T;\tau,\theta) = \int_0^T |A(s,\tau,\theta) - A(s,\tau,\theta_0)||z_1(s,\tau,\theta_0)| ds$. We have $A(t;\tau,\theta)$ is continuous due to assumptions on $G(x(t),\tilde{x}(t),\theta)$, so that as $\theta \to \theta_0$

$$\lim_{\theta\to\theta_0} |r(T;\tau,\theta)| = 0.$$

Applying Gronwall's inequality and taking the limit as $\theta \to \theta_0$ in (18), we have

$$\lim_{\theta\to\theta_0} |\Delta_3(t;\tau,\theta,\theta_0)| \le \lim_{\theta\to\theta_0} |r(T;\tau,\theta)| e^{Qt} = 0,$$

which yields $\lim_{\theta\to\theta_0} z_1(t,\tau,\theta) = z_1(t,\tau,\theta_0)$.

Next we argue that the solution to (17), depends continuously on the delay τ whenever $\dot{x}(t)$ is continuous. Let $\Delta_4(t;\tau,\tau^*,\theta) = z_1(t,\tau,\theta) - z_1(t,\tau^*,\theta)$ for a fixed $\theta \in \mathbb{R}^p$ and fixed $\tau^* \in [-r,0]$. We examine

$$
\begin{aligned}
|\Delta_4(t;\tau,\tau^*,\theta)| &= |z_1(t,\tau,\theta) - z_1(t,\tau^*,\theta)| \\
&\leq \int_0^t |A(s,\tau,\theta)z_1(s,\tau,\theta) - A(s,\tau^*,\theta_0)z_1(s,\tau^*,\theta)|ds, \\
&\leq \int_0^t \big\{|A(s,\tau,\theta)z_1(s,\tau,\theta) - A(s,\tau,\theta)z_1(s,\tau^*,\theta)| \\
&+ |A(s,\tau,\theta)z_1(s,\tau^*,\theta) - A(s,\tau^*,\theta)z_1(s,\tau^*,\theta)|\big\}ds, \\
&\leq \int_0^t |A(s,\tau,\theta)\Delta_4(s;\tau,\tau^*,\theta)|ds \\
&+ \int_0^t |(A(s,\tau,\theta) - A(s,\tau^*,\theta))z_1(s,\tau^*,\theta)|ds \\
&\leq Q\int_0^t |\Delta_4(s;\tau,\tau^*,\theta)|ds \\
&+ \int_0^t |[A(s,\tau,\theta) - A(s,\tau^*,\theta)]z_1(s,\tau^*,\theta)|ds \\
&\leq Q\int_0^t |\Delta_4(s;\tau,\tau^*,\theta)|ds + r^\tau(T;\tau,\theta),
\end{aligned} \tag{19}
$$

where

$$r^\tau(T;\tau,\theta) = \int_0^T |A(s,\tau,\theta) - A(s,\tau^*,\theta)||z_1(s,\tau^*,\theta)|ds.$$

Since $G_x,\ G_{\tilde{x}}$ are continuous and $x \in C^1(-\tau, T; \mathbb{R}^N)$, then as $\tau \to \tau^*$, $|A(t,\tau,\theta) - A(t,\tau^*,\theta)| \to 0$. Moreover, $A(t,\tau,\theta)$ is bounded in t,τ. As a result, when $\tau \to \tau^*$, the limit of $|r^\tau(T;\tau,\theta)|$ is 0. Applying Gronwall's inequality in (19) and taking the limit as $\tau \to \tau^*$, we find

$$\lim_{\tau\to\tau^*} |\Delta_4(t;\tau,\tau^*,\theta)| \leq \lim_{\tau\to\tau^*} |r^\tau(T;\tau,\theta)|e^{Qt} = 0,$$

and thus $\lim_{\tau\to\tau^*} z_1(t,\tau,\theta) = z_1(t,\tau^*,\theta)$. □

4. **Traditional and generalized sensitivity functions.** Having established the differentiability given above with respect to the quantities $\theta, z_0 = (x_0,\phi)$ and τ, we are now able to use powerful sensitivity techniques in analyzing delay systems. For further simplification in the remainder of our discussions we restrict our considerations to constant function initial conditions so in $z_0 = (x_0,\phi)$ we assume $\phi(\xi) = x_0,\ -\tau \leq \xi \leq 0$. Traditional sensitivity analysis is the quantification of the effect changes in parameters have on model solutions. Traditional sensitivity functions (TSFs), which are given by $y_1^k(t) = \dfrac{\partial x}{\partial \theta^k},\ k = 1,...,p,\ y_2^m(t) = \dfrac{\partial x}{\partial x_0^l},\ l = 1,...,n$, and $y_3(t) = \dfrac{\partial x}{\partial \tau}$, are local in nature. That is, they are defined by locally evaluated partial derivatives, i.e., $\dfrac{\partial x}{\partial \theta}(t,\bar{\theta},\bar{x}_0,\bar{\tau})$, which gives information over specified time intervals, and at values of parameters, initial conditions and delays. In spite of this locality in nature, these functions have been used to improve sampling in an experimental setting. In particular they can be used to guide the time at which measurements should be taken to best inform the estimation of unknown

parameters [16, 17]. Thus sampling might be advisable in time intervals where, for example, $y_1^k(t)$ is large, since this indicates that the model solution $x(t)$ is sensitive to changes in the parameter θ_k. Similarly, insensitivity to a certain parameter (or unknown quantity), as indicated by a small or near zero value of the TSF, implies that observations can not be profitably taken in that region if the goal is estimation of the parameter.

TSFs may be approximated by forward differences, but more typically are found by solving the system of sensitivity equations

$$\frac{d}{dt}\frac{\partial x(t)}{\partial \theta} = \frac{\partial G}{\partial x}\frac{\partial x}{\partial \theta}(t) + \frac{\partial G}{\partial \tilde{x}}\frac{\partial x}{\partial \theta}(t-\tau) + \frac{\partial G}{\partial \theta}(t) \tag{20}$$

along with the corresponding system

$$\begin{aligned} \frac{dx(t)}{dt} &= G(x(t), x(t-\tau), \theta), \quad t > 0 \\ x(\xi) &= x_0,\ -\tau \le \xi \le 0. \end{aligned} \tag{21}$$

Here the $\frac{\partial}{\partial \theta}$ and $\frac{d}{dt}$ operators have been interchanged, which is permissible due to the continuity assumptions made on G and x. We note that sensitivity analysis is most efficiently carried out in two steps. Once a solution $x(t)$ corresponding to $(\bar{\theta}, \bar{x}_0, \bar{\tau})$ of the above (original delay) equation (21) is obtained, one uses this solution to evaluate the coefficients in system (20). This decoupling of the original equation and the sensitivity equation has implications when considering the sensitivity with respect to the time delay τ (which if solved in a coupled manner would result in a neutral delay system, resulting in additional questions to be resolved).

Generalized sensitivity functions were first introduced by Thomaseth and Cobelli [66]. They were further studied in a series of papers by Banks, et al., [16, 17, 20], and provide a measure of how informative measurements of the output or observation variables ($f(t,q)$ defined below, which are not necessarily simply the state variables), are for the identification of unknown quantities. Here the functions G and h are assumed to be sufficiently differentiable to construct the TSFs as well as the generalized sensitivity functions (GSFs). Before defining the GSFs we briefly outline an inverse problem framework, not only to put our discussion in context, but also to define quantities in the definition of the GSFs.

Given a model solution $x(t)$, the (traditional) sensitivity of the solution with respect to an estimated quantity q_k (where $q = (\theta, x_0, \tau)^T$) is $ts_k(t,q) = \frac{\partial f}{\partial q_k}(t,q) \in \mathbb{R}^m$, where $f(t,q) = h(t, x(t), x(t-\tau), \theta)$ are the model quantities corresponding to the observed data. For comparison among parameters, (i.e., the solution is more sensitive to parameter q_i as compared to q_j for $i \neq j$), the relative sensitivity functions, given by $rs_k(t) = \frac{q_k}{f(t,q)}\frac{\partial f}{\partial q_k}(t,q)$, are sometimes computed.

Observations are typically available at discrete times, which we denote by $t_1, ..., t_{n_d}$. The model representation of the data is then $f(t_j, q) = h(t_j, x(t_j), x(t_j - \tau), \theta), \quad j = 1, \dots, n_d$. In general, the data are not exactly $f(t_j, q)$, due to uncertainty in the measurement process, and also due to small fluctuations not explicitly included in the model. Therefore we represent the observation process Y_j at time t_j by the *statistical model*

$$Y_j = f(t_j, q^0) + \mathcal{E}_j, \quad j = 1, \dots, n_d, \tag{22}$$

where $f(t_j, q) = h(t_j, x(t_j), x(t_j - \tau), \theta)$, $q = (\theta, x_0, \tau)$, for $q \in \mathcal{Q} = \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R}^1$. Here $q^0 = (\theta^0, x_0^0, \tau^0)^T$ represents the 'true values' of the parameters that generate the observations $\{Y_j\}_{j=1}^{n_d}$. The existence of q^0 is commonly assumed [15], implying that (8) is sufficient to describe the biological, sociological, or physical process precisely.

The observation errors $\mathcal{E}_j$ are assumed to be random variables with unknown but assumed independent and identical probability distributions of mean zero and constant variance σ^2. Each data set $\{y_j\}_{j=1}^{n_d}$ is one realization of the random variable $\{Y_j\}_{j=1}^{n_d}$, and the corresponding errors ϵ_j are also realizations of the $\mathcal{E}_j$. Assuming the statistical model (22) and estimating unknown quantities via the minimization between the model and data gives rise to the commonly used ordinary least squares (OLS) estimator defined by

$$q_{OLS} = \arg\min_{q \in \mathcal{Q}} \sum_{j=1}^{n_d} |Y_j - f(t_j, q)|^2 . \tag{23}$$

Here the objective functional is minimized over an admissible parameter space $\mathcal{Q}$. Another common formulation assumes relative error in a weighted least squares procedure. In this case the error is assumed to be proportional to the model quantity $f(t_j; q)$.

The variance σ^2 of the observation error is used in the computation of standard errors, confidence intervals, etc., and also in the generalized sensitivity functions. For a given set of data, $\{y_j\}_{j=1}^{n_d}$ and parameter estimates $\hat{q}$, the (bias-adjusted) variance is estimated as

$$\hat{\sigma}^2 = \frac{1}{n_d - n_p} \sum_{j=1}^{n_d} |y_j - f(t_j, \hat{q})|^2 \tag{24}$$

for $n_p = p + n + 1$ estimated parameters, where $n_p = dim(\mathcal{Q})$. Under reasonable smoothness and regularity assumptions, one can use arguments from (nonlinear) asymptotic theory to argue that as $n_d \to \infty$, $q_{OLS} \sim \mathcal{N}_{n_p}(q^0, \Sigma_0^{n_d}) = (q^0, \sigma^2[\chi(q^0)^T \chi(q^0)]^{-1})$. The $n_d \times n_p$ sensitivity matrix $\chi(q)$ is made up of elements

$$\chi_{jk} = \frac{\partial f(t_j, q)}{\partial q_k} \quad \text{for } j = 1, ..., n_d, k = 1, ..., n_p.$$

Then the standard error for each estimate $\hat{q}_k$ is computed as $SE(\hat{q}_k) = \sqrt{\hat{\Sigma}_{kk}}$ where the covariance matrix Σ is estimated as $\Sigma \approx \hat{\Sigma} = \hat{\sigma}^2[\chi(\hat{q})^T \chi(\hat{q})]^{-1}$. For a more complete discussion of the underlying assumptions and related formulations, the reader may see [15].

The generalized sensitivity functions [17, 20, 66] are defined by

$$\mathbf{gs}(t) = \int_0^t \left[F(T)^{-1} \frac{1}{\sigma^2(s)} \nabla_q f(s, q^0) \right] \cdot \nabla_q f(s, q^0) dP(s), \quad t \in [0, T], \tag{25}$$

for variance $\sigma^2(t)$ that may possibly be time-dependent, true parameters q^0, some general measure P that embodies the observations, and the Fisher information matrix (FIM) F which is defined by

$$F(T) = \int_0^T \frac{1}{\sigma^2(t)} \nabla_q f(t, q^0) \nabla_q f(t, q^0)^T dP(t). \tag{26}$$

We note that the definition of the measure P affects the FIM, and one goal would be to choose it in such a way as to optimize the information from data concerning the estimated parameters. The GSFs are cumulative functions, such that at time t_j, only the contributions of measurements up to and including those at time t_j are relevant. By the definition in (25), it is readily seen that the GSFs are one at the final time $\mathbf{gs}(t_{n_d}) = 1$. As discussed in [17, 66], regions over which the *sharpest change* (either increase or decrease) of the GSFs indicate regions of high information content. Decreases in the GSF corresponding to a given parameter indicate correlation between that parameter and at least one other estimated parameter. In this case, it can be seen [17] that computing the GSF for one of the correlated parameters and holding the others fixed, will result in a monotonically increasing GSF. Therefore, regions over which the GSF decreases indicate that the data in that region indeed contains information concerning that parameter, but it is correlated with at least one other parameter, and simultaneous identifiability of all parameters may be difficult.

As observations are typically available at discrete time points and our discussions are in the context of parameter estimation from observed or measured data, we have included here also the definitions for the GSFs and FIM for a discrete measure $P = \sum_{j=1}^{n_d} \Delta_{t_j}$. In the discrete case, the generalized sensitivity functions are

$$\mathbf{gs}(t_j) = \sum_{i=1}^{j} \frac{1}{\sigma^2(t_j)} \left[F^{-1} \times \nabla_q f(t_i, q^0)\right] \cdot \nabla_q f(t_i, q^0), \tag{27}$$

for observation times t_j where $j = 1, ..., n_d$. In the above definition, the discrete FIM is given by

$$F = \sum_{j=1}^{n_d} \frac{1}{\sigma^2(t_j)} \nabla_q f(t_j, q^0) \nabla_q f(t_j, q^0)^T, \tag{28}$$

which measures the information content of the data corresponding to the parameters. In both (25) and (27), the (biased) estimate for the variance of the observation error is used up to and including the time t_j of the observation, given by

$$\sigma^2(t_j) = \frac{1}{j} \sum_{i=1}^{j} \left|y_i - f(t_i, \hat{q})\right|^2. \tag{29}$$

If the variance is assumed constant ($\sigma^2(t) \equiv \sigma^2$), one would simply calculate the estimate as in (24), and use that in (25) or (27).

5. **Computational examples.**

We illustrate the uses of sensitivity analysis in two prominent examples of delay equations. The first example we consider, the logistic equation, is a delay version of one of the most commonly studied models of growth/decay. This delayed logistic equation, commonly referred to as Hutchinson's equation, is not only discussed in most introductory modeling courses, but is still used in research endeavors to represent growth within an environment in which saturation is possible, but the death rate is proportional to previous population levels. The standard (without delay) logistic example has been effectively used to illustrate with simulated data the ideas of traditional and sensitivity functions and how these techniques may improve data sampling for the purpose of parameter estimation [16, 17]. It is thus quite natural to turn to the delayed logistic equation now that we are able to study sensitivity functions in systems involving a discrete delay. We will numerically generate simulated data with a known delay, and demonstrate that the estimation can be improved using insights gained from the sensitivity function solutions.

The second example we use is also a standard model, the delayed harmonic oscillator of Minorsky discussed in the Introduction. As noted there, this example arises in many physical applications where oscillatory phenomena are important and either delayed restoring or delayed damping is relevant.

5.1. **Hutchinson equation example.** In his by now classical paper [50] (and in [51]), Hutchinson arrived at a version of the logistic equation that incorporated a delay in the carrying capacity or death rate term given by

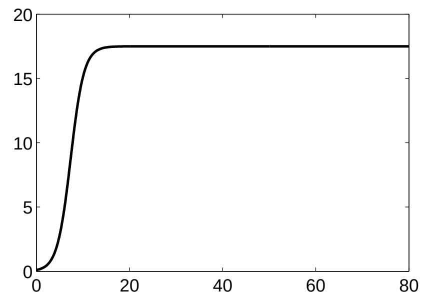
$$\frac{dx(t)}{dt} = rx(t)\left(1 - \frac{x(t-\tau)}{K}\right). \tag{30}$$

The model was suggested as a possible explanation of the growth dynamics seen in Daphnia. This population seemed to grow exponentially at low population sizes, but it would oscillate at higher population levels. Hutchinson hypothesized that this growth was like that of the logistic model, only that the population seemed to be able to exceed its carrying capacity and perhaps it was around this value that the population level was oscillating.

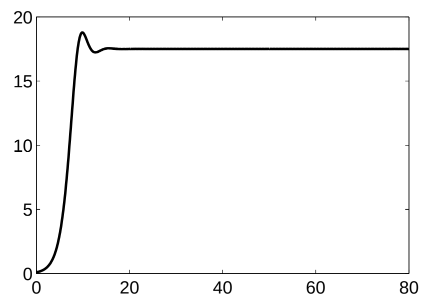
The model, given as a footnote in the original paper, can be interpreted as the population growing essentially exponentially at low population sizes, just as with the traditional logistic growth model. The delay in the second term results in the population being able to exceed the carrying capacity upon initial growth, since as the size of the population increases to the carrying capacity K, it is the population size at a previous time $(t-\tau)$ that is less than the carrying capacity K, that determines the growth rate. Thus, as $x(t)$ increases to $x(t) = K$, $\frac{dx}{dt} > 0$ and the population continues to grow. The population size continues to increase until the population size at $t-\tau$ reaches the carrying capacity $(x(t-\tau) \geq K)$, when the population will decrease. Similarly, due to this delay, as time continues the population will decrease below the carrying capacity since as $x(t)$ reaches the carrying capacity K, $x(t-\tau) \geq K$ and thus $\frac{dx}{dt} < 0$. The population size will continue to oscillate (perhaps with some damping) in this fashion around its carrying capacity K. The effect of the delay can be seen in Figure 1, in which solutions for the Hutchinson equation without delay, and for a small and moderate delay are shown.

Hutchinson, a zoologist, was interested in a topic in ecology which he termed 'circular causal systems'. These are essentially systems of interacting populations, or a system composed of a population or populations which affect their environment, and the changed environment then affects the population. He postulated, without specifically naming them, that feedback loops are possible: that if a set of properties in either system changes in such a way that the action of the first system on the second changes, this may cause changes in properties of the second system, which alter the mode of action of the second system on the first. He also considered that observed oscillations may be a result of competition, and therefore better described by competition models, depending on the nature of the feedback, or interaction between the two interacting systems.

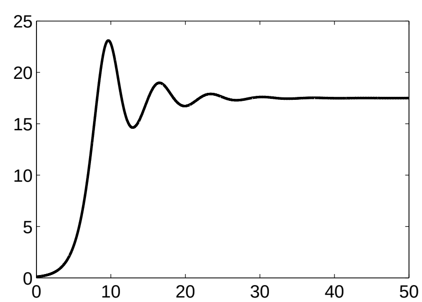
He viewed carrying capacity as a self-regulating mechanism, and speculated that it wasn't exceeded often, because to do so, would disrupt a delicate balance in



(A) $\tau = 0$



(B) $\tau = 1$



(C) $\tau = 1.5$

FIGURE 1. Solutions to the Hutchinson equation without delay (a), with small delay $\tau = 1$ (b), and with moderate delay $\tau = 1.5$ (b). The initial condition used is $x_0 = 0.1$ and parameter values are $r = 0.7$ and $K = 17.5$. These values will be used in all other discussed computations unless otherwise noted.

nature, highly destabilizing the population, potentially through the introduction of new unfavorable conditions, and thereby subjecting it to greater risk of extinction. Therefore, he surmised that it is highly likely that there may be a general tendency for the time lag to be minimized by natural selection, and that the environment would strongly influence the population size to be reduced at or below the carrying capacity.

In an alternative formulation the delay could appear in other terms, and does in the version of the delay logistic studied by E. M. Wright [67]. This version is

$$y'(t+1) = -ry(t)\left\{1 + y(t+1)\right\},$$

which can be obtained from equation (30) by a change of variables (e.g., set $y(x(t)) = -1 + \frac{x(t)}{K}$). As such, the delay logistic is sometimes called the Hutchinson-Wright Equation. Wright referred to this equation as a difference-differential equation, and worked toward developing a general theory of these classes of equations in a series of papers. The delay logistic as in (30) has remained a well-studied example of a classic model with dramatic qualitative behavior changes due to the inclusion of a delayed effect.

Concerning Daphnia, Hutchinson cited a paper by Pratt in 1946 [61] and noted that such a delay could be interpreted as the observed fertility of a parthenogenetic female (capable of reproduction from unfertilized ovum) impacted by population density at current time and also at previous times to which the ovum has been exposed. He also cited the dynamics in Elton (1942) [37] and Errington (1946) [38] as other possible examples of such dynamics in their studies of rodent populations. In these populations it was observed that intraspecies fighting increased at large population densities, especially among younger animals about to embark on their reproductive life stages. Then, when a population nears its carrying capacity, the growth rate may be influenced by the population density at a time in the past, when the size of the population τ units in the past is less than at the present time, $N(t-\tau) < N(t)$, and the population thus increases as the sexually mature/active animals were those who experienced less intraspecies fighting. Therefore, even at carrying capacity, the population can grow, exceeding this level until the juveniles who experienced increased fighting reach reproductive age.

The traditional sensitivity functions with respect to the model parameters r, K, initial condition x_0, and delay τ are given by

$$\frac{\partial ts_r(t)}{\partial t} = r\left[1 - \frac{x(t-\tau)}{K}\right] ts_r(t) - \frac{rx(t)}{K} ts_r(t-\tau) + x(t)\left[1 - \frac{x(t-\tau)}{K}\right] \quad (31)$$

$$\frac{\partial ts_K(t)}{\partial t} = r\left[1 - \frac{x(t-\tau)}{K}\right] ts_K(t) - \frac{rx(t)}{K} ts_K(t-\tau) + rx(t)\left[\frac{x(t-\tau)}{K^2}\right] \quad (32)$$

$$\frac{\partial ts_{x_0}(t)}{\partial t} = r\left[1 - \frac{x(t-\tau)}{K}\right] ts_{x_0}(t) - \frac{rx(t)}{K} ts_{x_0}(t-\tau) \quad (33)$$

$$\frac{\partial ts_\tau(t)}{\partial t} = r\left[1 - \frac{x(t-\tau)}{K}\right] ts_\tau(t) - \frac{rx(t)}{K}\left[ts_\tau(t-\tau) - \dot{x}(t-\tau)\right], \quad (34)$$

where $ts_r(t) = \frac{\partial x(t)}{\partial r}$, $ts_K(t) = \frac{\partial x(t)}{\partial K}$, $ts_{x_0}(t) = \frac{\partial x(t)}{\partial x_0}$, and $ts_\tau(t) = \frac{\partial x(t)}{\partial \tau}$. As noted earlier, we consider only the case of constant initial data, and thus we do not discuss here the Fréchet derivative $y_2(t) = \frac{\partial}{\partial z_0} x(t, z_0, \theta)$ where $z_0 = (x_0, \phi), Z = \mathbb{R}^n \times L^2(\tau, 0; \mathbb{R}^n)$; the results of Theorem 6 still ensure the existence and uniqueness of the solution $\frac{\partial x(t)}{\partial x_0}$ to equation (33), for this simpler case. The existence of unique solutions to equations (31) and (32) are guaranteed by Theorem 5, and a unique solution for equation (34) by Theorem 7. Note that equation (34) is not a neutral equation if one assumes the solution $x(t)$ (and also $x(t-\tau)$) is already computed when sensitivity analysis is done. Thus we decouple the original equation and first solve the delay equation before computing sensitivities. Here when computing sensitivities the $x(t)$ and $x(t-\tau)$ are not unknown quantities but rather an input in the traditional sensitivity functions above.

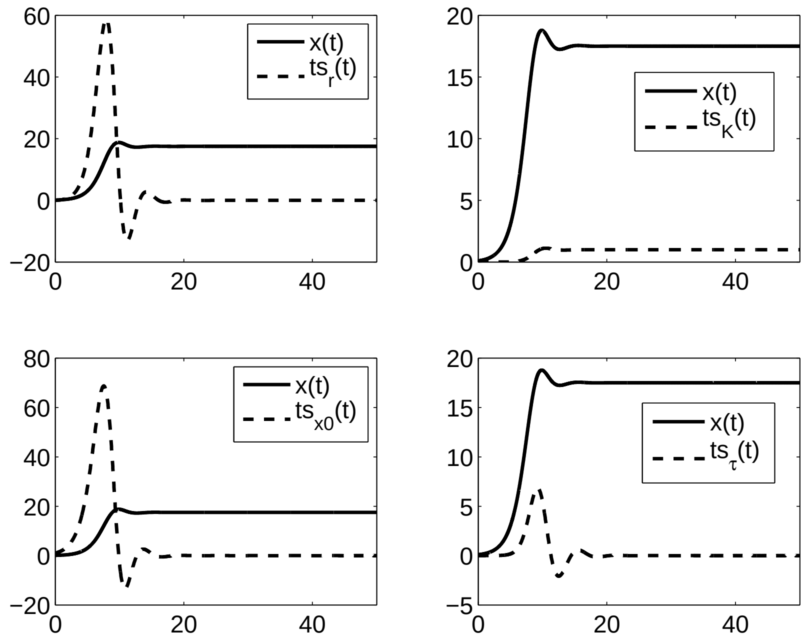


FIGURE 2. The solution for the Hutchinson equation overlayed with the traditional sensitivity functions with respect to growth rate r, carrying capacity K, constant initial state x_0, and delay τ, each evaluated at $(r, K, x_0, \tau) = (0.7, 17.5, 0.1, 1)$.

The solutions for the Hutchinson equation with small delay $\tau = 1$ overlayed with the traditional sensitivity functions $\mathbf{ts}(t)$ can be found in Figure 2. As one would expect, the solution is most sensitive to the initial condition for early times, and slightly lagging the $ts_{x_0}(t)$ trajectory is the sensitivity with respect to the growth rate r. The solution is sensitive to the carrying capacity when it is near and after it has reached it. All of these characteristics are consistent with the logistic equation, which is the Hutchinson equation without delay ($\tau = 0$). The solution appears sensitive to the delay, that is $ts_\tau(t)$ is elevated for a short period corresponding to the same interval ($t \in [7.5, 14]$) over which the solution first exceeds its carrying capacity and then decreases below it, and finally approaches it around $t \approx 14$. It makes sense that the solution has increased sensitivity to the delay τ in the same region that we see the difference between including the delay and no delay. While one would be tempted to conclude, from the magnitudes of the peaks reached by the traditional sensitivity functions shown in Figure 2, that the

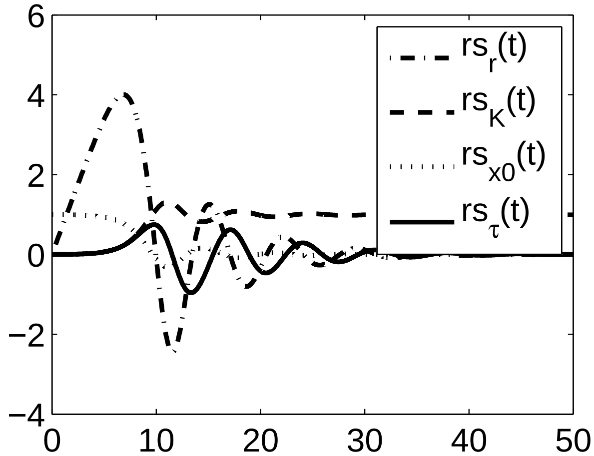


FIGURE 3. The relative sensitivity functions for the growth rate r, carrying capacity K, initial state x_0 and delay τ for the Hutchinson equation corresponding to the traditional sensitivity functions as in Figure 2.

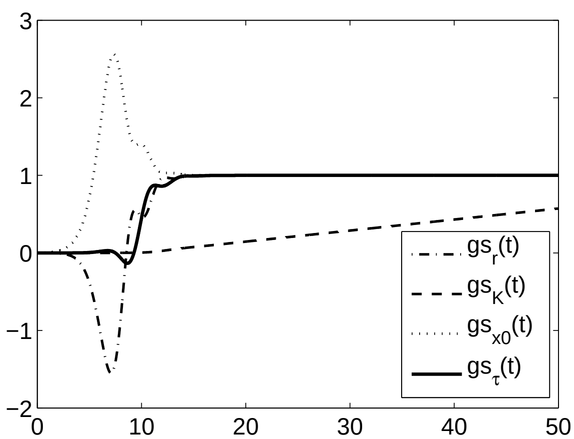


FIGURE 4. The generalized sensitivity functions for the growth rate r, carrying capacity K, initial state x_0 and delay τ for the Hutchinson equation corresponding to the traditional sensitivity functions as in Figure 2.

solution is most sensitive to the initial condition x_0, the relative sensitivity functions suggest otherwise. The relative sensitivity functions, which can be used to compare sensitivity among different parameters, corresponding to the Hutchinson equation for these values can be found in Figure 3. The generalized sensitivity functions can be found in Figure 4, and support that the regions of sensitivity are as suggested by the traditional and relative sensitivity functions. The solutions for $gs_r(t)$ and

$gs_{x_0}(t)$ suggest that these quantities are correlated during the time intervals in which the solution is sensitive to each, since an increase in one curve occurs along with a simultaneous decrease in the other. Thus, estimating both the initial condition x_0 and the growth rate r from data corresponding to this interval is likely problematic.

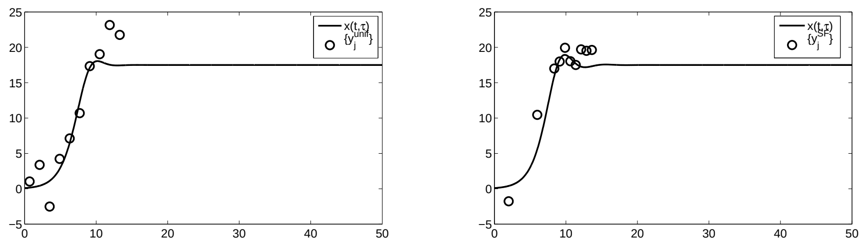
To illustrate the information gained from the solutions of the sensitivity functions with a delay $\tau = 1$, we generated simulated data according to statistical model (22) with 10% error. We then performed least squares estimation (via equation (23)) to estimate the delay τ, while holding the other parameters fixed. As seen in [17], any parameter correlation issues are then irrelevant and estimates should be improved if data is concentrated in any regions of enhanced information content as suggested by high or low values of the $ts_\tau(t)$, and regions of greatest change in $gs_\tau(t)$.

Estimates of the delay τ are contained in Table 1, from data $\{y_j^{unif}\}_{j=1}^{10}$ uniformly sampled over the interval $t \in [0, 14]$, and data $\{y_j^{SF}\}_{j=1}^{10}$ in which 8 of the 10 points are concentrated in the region $t \in [7.5, 14]$ as informed by the sensitivity functions. The estimations were done with initial values for $\hat{\tau}$ both 20% above and below the true value of $\tau = 1$. The estimated delays $\hat{\tau}$ from the data concentrated in the region of sensitivity to the delay is considerably closer than when uniform data is used. Additionally, the lower standard errors confirm that we should have more confidence, or rather that the estimate is more reliable, in $\hat{\tau}$ estimated from $\{y_j^{SF}\}_{j=1}^{10}$. The solutions of the Hutchinson equation with the estimated values $\hat{\tau}$ overlayed with the data from which the delays were estimated are shown in Figure 5. The improvement in the agreement between model solution and data is apparent in comparisons of Figures 5a and 5b.

TABLE 1. Estimation of delay τ, from data $\{y\}$ generated with true value $\tau = 1$.

	initial $\hat{\tau}$	$\hat{\tau}$	$SE(\hat{\tau})$
$\{y_j^{unif}\}_{j=1}^{10}$	0.8	0.7862	0.14
$\{y_j^{SF}\}_{j=1}^{10}$	0.8	1.0247	0.09
$\{y_j^{unif}\}_{j=1}^{10}$	1.2	0.8525	0.13
$\{y_j^{SF}\}_{j=1}^{10}$	1.2	1.0247	0.09

Further least squares estimation computations were done to establish that estimation of the delay along with other parameters are possible, although results are not shown. Not surprisingly, estimating any of the quantities r, K, and x_0 as long as sufficient data in each region of sensitivity of the quantity to be estimated is present, is possible within a reasonable degree of accuracy. We would only expect difficulty in the simultaneous estimation of the delay with other quantities if correlation appears to be a problem, possibly apparent in the generalized sensitivity functions. However, while there is a decrease in $gs_\tau(t)$ over a small time interval, it is not sufficient to result in difficulty estimating this parameter along with any other model quantitiy. Also, the increases in $gs_\tau(t)$ elsewhere do not seem to correspond to decreases in generalized sensitivity functions for other model quantities. Therefore, from the sensitivity functions, we have no reason to suspect correlation in the estimation of the delay with other model quantities to the extent that estimation efforts would be difficult, and computations (not included here) have confirmed that expectation.



(A) 'Best fit' solution from $\{y_j^{unif}\}_{j=1}^{10}$

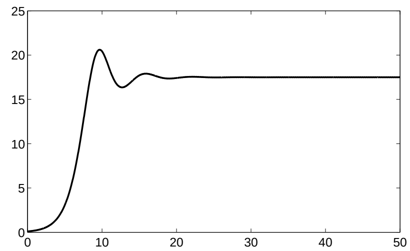
(B) 'Best fit' solution from $\{y_j^{SF}\}_{j=1}^{10}$

FIGURE 5. The solutions to the delay logistic equation with estimated delay $\hat{\tau}$ from data as shown in each graph. In (a) the solution is shown with estimated $\hat{\tau}$ from 10 data points spread uniformly over the time interval $[0, 14]$. In (b) the solution is shown with estimated $\hat{\tau}$ from 8 of the total 10 data points concentrated in the time interval $[7.5, 14]$.

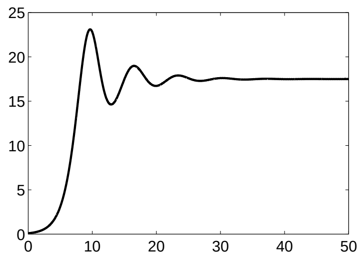
We demonstrate similar findings as the delay is increased to $\tau = 1.25$ and $\tau = 1.5$. Solutions to the Hutchinson equation with these delays along with corresponding traditional and generalized sensitivity functions are shown in Figure 6. The solutions (in 6a and 6b) oscillate for increasingly long time periods. The sensitivity functions show the expected trends for parameters r, K, and x_0 (as well as the inverse relationships between $gs_r(t)$ and $gs_{x_0}(t)$). However, the oscillations for the sensitivity functions, particularly oscillations in $ts_\tau(t)$ and increases in $gs_\tau(t)$ persist when oscillations in the solution $x(t)$ are barely discernible. In this case, the insight gained from the sensitivity functions indicates when one can expect data to be informative about model parameters and the delay if taken from time intervals in which it appears that the system has reached its steady state and no observable dynamics are apparent.

Estimation results from data concentrated in the region suggested from the sensitivity functions $\{y^{SF}\}$ and uniformly distributed $\{y^{unif}\}$ over $t \in [0, 50]$, are contained in Table 2, and examples of solutions from estimated data are shown in Figure 7. Due to the region of sensitivity for the larger delay $\tau = 1.5$ being relatively long, fewer points were used to demonstrate the advantage of data concentrated in the region of sensitivity and uniform data. That is, if more total data points in each set were used, many of them in the uniform data set were within the region of sensitivity. The results in the left side of Table 2 are from the estimation of true delay $\tau = 1.25$, and suggest a marked advantage of using data $\{y^{SF}\}$ as compared with $\{y^{unif}\}$ observations. The estimates $\hat{\tau}$ are closer to the true value and the comparably smaller standard error when an initial guess 20% below the true value (initial $\hat{\tau} = 1$) suggests that it is also more reliable.

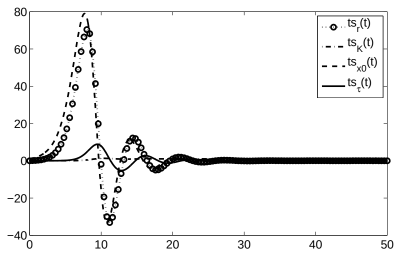
The lower standard error for the estimate obtained using uniform data when the initial guess $\hat{\tau} = 1.5$ is 20% above the true value can be explained by observing that the solution with the estimated delay $\hat{\tau} = 1.674$ (Figure 7a) is very near the data points, resulting in a small objective functional. This leads to a small 'estimated'



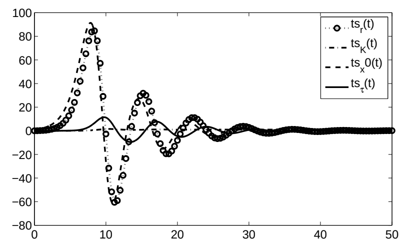
(A) solution $x(t)$ with $\tau = 1.25$



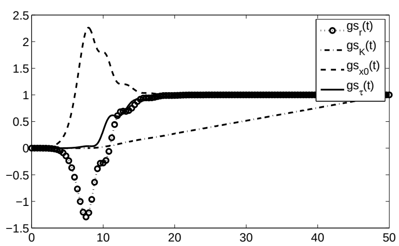
(B) solution $x(t)$ with $\tau = 1.5$



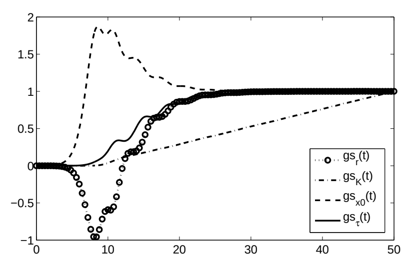
(C) TSF with $\tau = 1.25$



(D) TSF with $\tau = 1.5$



(E) GSF with $\tau = 1.25$



(F) GSF with $\tau = 1.5$

FIGURE 6. Solutions to the Hutchinson equation for delay $\tau = 1.25$ (a) and $\tau = 1.5$ (b). The traditional sensitivity functions $\mathbf{ts}(t)$ with respect to r, K, x_0 and τ are shown in (c) for delay $\tau = 1.25$ and (d) for delay $\tau = 1.5$. Corresponding generalized sensitivity functions $\mathbf{gs}(t)$ are shown in (e) and (f) for $\tau = 1.25$ and $\tau = 1.5$, respectively.

$\hat{\sigma}^2$, computed from equation (29), as one would do in a typical estimation procedure. The sensitivity matrix $\chi(\hat{\tau})$ is still relatively rank-deficient (a consequence of using multiple data points with similar sensitivity to the estimated parameters, as discussed in [16]) as compared with the sensitivity matrix from the observations $\{y^{SF}\}$, and the resulting small standard error is therefore misleading, and is not obtained when using other initial guesses for $\hat{\tau}$ (not shown). We caution against being overly confident in estimation results beginning from only one initial estimate for the estimated parameters.

TABLE 2. Estimation of delay τ, from simulated data $\{y\}$. For 'true' delay $\tau = 1.25$, six data points ($n_d = 6$) were used, with $\{y_j\}_{j=1}^{5}$ concentrated in the interval $t \in [8, 16]$. For 'true' delay $\tau = 1.5$, five data points ($n_d = 6$) were used, with $\{y_j\}_{j=1}^{5}$ concentrated in the interval $t \in [8, 22]$. For delay $\tau = 1.25$, $\{y_j^{unif}\}_{j=1}^{6}$ was spread uniformly over the time interval $[0, 50]$ and $\tau = 1.5$,

	'true' τ	init $\hat{\tau}$	$\hat{\tau}$	$SE(\hat{\tau})$
$\{y_j^{unif}\}_{j=1}^{5}$	1.25	1	1.000	0.5
$\{y_j^{SF}\}_{j=1}^{6}$	1.25	1	1.31	0.08
$\{y_j^{unif}\}_{j=1}^{5}$	1.25	1.5	1.674	0.07
$\{y_j^{SF}\}_{j=1}^{6}$	1.25	1.5	1.31	0.08
$\{y_j^{unif}\}_{j=1}^{5}$	1.5	1.2	1.201	1
$\{y_j^{SF}\}_{j=1}^{6}$	1.5	1.2	1.523	0.06
$\{y_j^{unif}\}_{j=1}^{5}$	1.5	1.8	1.582	0.1
$\{y_j^{SF}\}_{j=1}^{6}$	1.5	1.8	1.523	0.06

5.2. **Harmonic oscillator.** We turn finally to illustrating the use of the TSF and GSF for the Minorsky harmonic oscillators with delays as described in the Introduction. We recall that the equation with delayed damping has the form

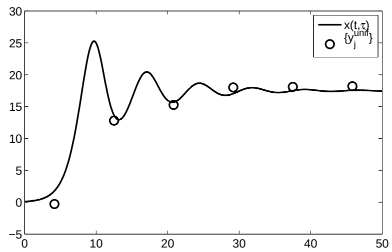
$$\frac{d^2x(t)}{dt^2} + K\frac{dx(t-\tau)}{dt} + bx(t) = g(t), \tag{35}$$

while the system with delayed restoring force is given by

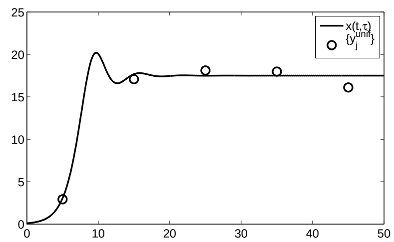
$$\frac{d^2x(t)}{dt^2} + K\frac{dx}{dt} + bx(t-\tau) = g(t). \tag{36}$$

We use traditional and generalized sensitivity functions with equations (35) and (36) and illustrate their application in determining regions of sensitivity for model parameters K, b and time delay τ. As before, we take the derivative of equation (35) with respect to each parameter q_i, where $q = (K, b, \tau)^T$ to obtain the TSF corresponding to that parameter q_i. First, letting $x = x_1(t)$ and $x_2(t) = \dot{x}(t)$, and rewriting equation (35) as a first order system we have

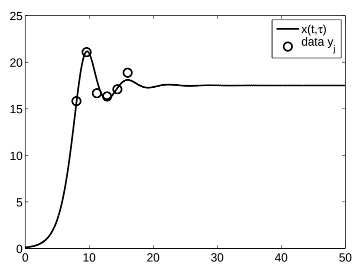
$$\begin{aligned}\frac{dx_1(t)}{dt} &= x_2(t),\\ \frac{dx_2(t)}{dt} &= g(t) - bx_1(t) - Kx_2(t-\tau).\end{aligned} \tag{37}$$



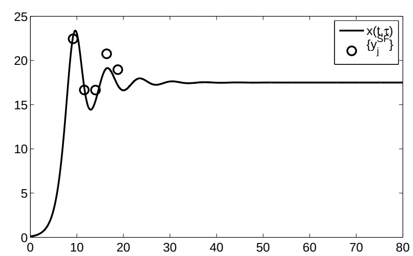
(A) 'Best fit' from $\{y_j^{unif}\}_{j=1}^{6}$; 'true' $\tau = 1.25$



(B) 'Best fit' from $\{y_j^{unif}\}_{j=1}^{5}$; 'true' $\tau = 1.5$



(C) 'Best fit' from $\{y_j^{SF}\}_{j=1}^{6}$; 'true' $\tau = 1.25$

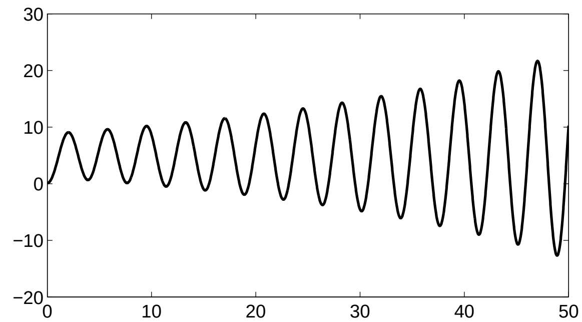


(D) 'Best fit' from $\{y_j^{SF}\}_{j=1}^{5}$; 'true' $\tau = 1.5$

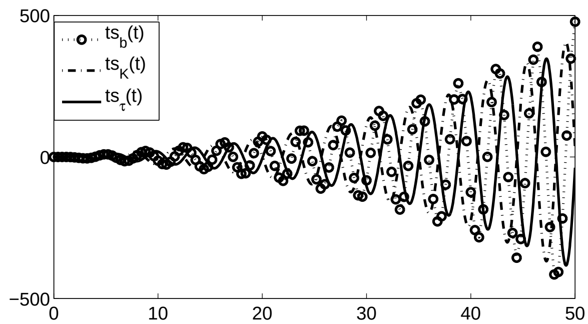
FIGURE 7. The solutions to the delay logistic equation with estimated delay $\hat{\tau}$ from data as shown in each graph. Panels (a) and (c) are the resulting estimates from data generated with $\tau = 1.25$ from $\{y^{unif}\}$ observations and $\{y^{SF}\}$ observations, respectively. Panels (b) and (d) are the resulting estimates from data generated with $\tau = 1.5$ from $\{y^{unif}\}$ observations and $\{y^{SF}\}$ observations, respectively. Further description of those data sets is in the caption of Table 2.

The traditional sensitivity functions are then solutions of

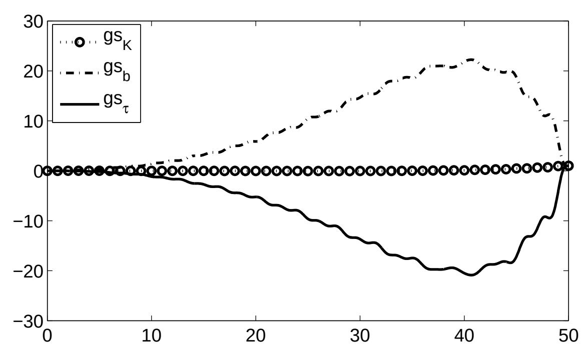
$$\begin{aligned}
\frac{ds_1(t)}{dt} &= s_4(t),\\
\frac{ds_2(t)}{dt} &= s_5(t),\\
\frac{ds_3(t)}{dt} &= s_6(t),\\
\frac{ds_4(t)}{dt} &= -bs_1(t) - Ks_4(t-\tau) - x_2(t-\tau),\\
\frac{ds_5(t)}{dt} &= -bs_2(t) - Ks_5(t-\tau) - x_1(t),\\
\frac{ds_6(t)}{dt} &= -bs_3(t) - Ks_6(t-\tau) + K\dot{x}_2(t-\tau),
\end{aligned}$$



(A) solution $x(t)$



(B) traditional sensitivity functions $ts(t)$



(C) generalized sensitivity functions $gs(t)$

FIGURE 8. Depicted above are (a) the solution to the harmonic oscillator with delayed damping $K = .5, b = 2$, $\tau = 1$, and $g(t) = 10$, (b) the traditional sensitivity functions and (c) the generalized sensitivity functions with respect to K, b, τ.

for $s_1(t) = \dfrac{\partial x_1(t)}{\partial K}$, $s_2(t) = \dfrac{\partial x_1(t)}{\partial b}$, $s_3(t) = \dfrac{\partial x_1(t)}{\partial \tau}$, $s_4(t) = \dfrac{\partial x_2(t)}{\partial K}$, $s_5(t) = \dfrac{\partial x_2(t)}{\partial b}$, and $s_6(t) = \dfrac{\partial x_2(t)}{\partial \tau}$.

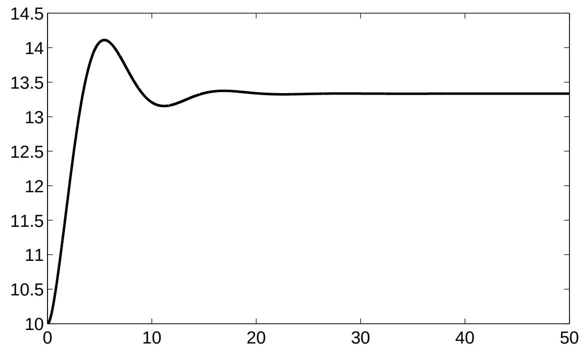
In Figure 8, the solution for the harmonic oscillator with delayed damping is shown for parameter values $K = 0.5, b = 2, g(t) \equiv 10$, and delay $\tau = 1$, along with the solutions of the traditional and generalized sensitivity functions with respect to $q = (K, b, \tau)^T$. The solutions of the TSFs imply that the solution is sensitive to all three parameters, with the sensitivities varying out of phase with each other indicating the region of the oscillations which are most sensitive to each parameter. This is informative as data taken exactly in phase with the oscillator would be limited in its information content concerning at least one of the three parameters. Both the traditional and generalized sensitivity functions suggest that data taken in the beginning time intervals should not be expected to contain as much information about the parameters as when the oscillations are growing in amplitude. One should not conclude that small oscillations do not contain information about the parameters, but rather that large oscillations contain more information in comparison.

Immediately, one should be wary of possible correlation among these three quantities as their regions of sensitivity as shown in Figure 8b are identical. The solutions to the generalized sensitivity functions, however, clarify this point, and indicate that it is the parameter b and delay τ that are correlated and the parameter K is uncorrelated with the other two over its regions of sensitivity. The delay has been increased sufficiently for these parameter values so that it counteracts the damping effect. As the delay is increased from zero, the frequency and amplitude of the oscillations increase. The amplitude of the oscillations grow, rather than decay, for this value of the delay. The delay has a similar effect as increasing the restoring coefficient b, which also increases the frequency of the oscillations. Therefore, if one were to estimate parameters with this model, one should not expect to estimate both b and τ simultaneously, but estimating either b or τ does not affect one's ability to estimate the parameter K.

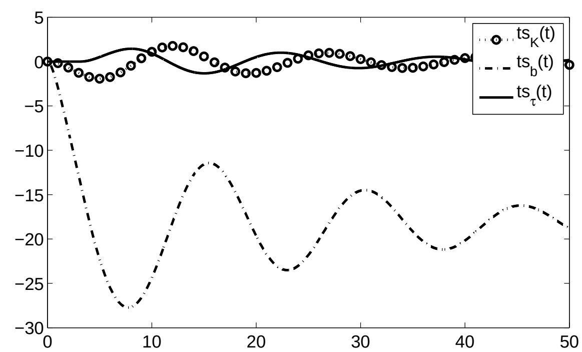
The sensitivity functions for the harmonic oscillator with delayed restoring force, equation (36), are arrived at in the same manner as when the delay appears in the damping term and are therefore omitted. The solution $x(t)$ and the corresponding traditional and generalized sensitivity solutions with respect to $q = (K, b, \tau)^T$ are shown in Figure 9. The solution, for this parameter range, appears similar to the Hutchinson equation with $\tau = 1$ in Figure 1b, in that it appears to exceed, and then drop below its steady state once before reaching it. The sensitivity functions, however, show much different behavior in that it appears that the sensitivity with respect to both parameters K and b and delay τ is nontrivial throughout much of the solution. It appears that the effects of each of these quantities (and the relative contributions of their terms) counteract each other to result in the observed steady state. While the sensitivity functions (more easily seen in $\mathbf{gs}(t)$ in Figure 9c) after around $t = 30$ do suggest that not much more information is to be gained with respect to K and τ, their regions of sensitivity again are much longer than what one would expect without the the insight the sensitivity functions provide.

6. Summary and final remarks.

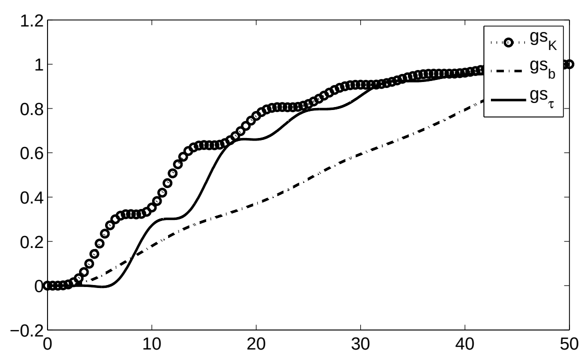
As evidenced by more than 50 years of literature, delay equations continue to play an important role in many areas of modeling, but especially in biological modeling. These include numerous areas of more classical as well as modern applications ranging from cellular biology to ecological systems to modern hereditary phenomenon in psychology and behavior theory. Of special interest are applications to experimental design and in particular sensitivity analysis and associated areas of uncertainty quantification in inverse problems.



(A) solution $x(t)$



(B) traditional sensitivity function $ts(t)$



(C) generalized sensitivity function $gs(t)$

FIGURE 9. Shown above are (a) the solution to the harmonic oscillator with delayed restoring force with $K = 2, b = 0.75$, $\tau = 3$, and $g(t) = 10$, (b) the traditional sensitivity functions and (c) the generalized sensitivity functions with respect to K, b, τ.

In the treatment above we have presented fundamental results guaranteeing that derivatives with respect delays, initial conditions and finite dimensional parameters exist, thereby justifying the formulation of traditional as well as generalized sensitivity functions. We have illustrated with classical examples how these sensitivities can be used to assist in design of inverse problems and, subsequently, control problems of interest in a wide variety of applications.

Acknowledgments. This research was supported in part by Grant Number NIAID R01AI071915-09 from the National Institute of Allergy and Infectious Disease, in part by the U.S. Air Force Office of Scientific Research under grant number FA9550-12-1-0188 and in part by the National Institute on Alcohol Abuse and Alcoholism under a subcontract from the Research Foundation for Mental Hygiene. The authors are grateful to several anonymous referees for their valuable comments and suggestions for improving this manuscript.

REFERENCES

[1] J. Arino, L. Wang and G. Wolkowicz, *An alternative formulation for a delayed logistic equation*, J. Theo. Bio., **241** (2006), 109–118.

[2] C. Baker and F. Rihan, *Sensitivity analysis of parameters in modelling with delay-differential equations,* MCCM Tec. Rep., **349** (1999).

[3] H. T. Banks, *Necessary conditions for control problems with variable time lags*, SIAM J. Control, **6** (1968), 9–47.

[4] H. T. Banks, *Representations for solutions of linear functional differential equations*, J. Differential Equations, **5** (1969), 399–409.

[5] H. T. Banks, *Delay systems in biological models: Approximation techniques*, in "Nonlinear Systems and Applications" (ed. V. Lakshmikantham), Academic Press, New York, (1977), 21–38.

[6] H. T. Banks, *Approximation of nonlinear functional differential equation control systems*, J. Optimiz. Theory Appl., **29** (1979), 383–408.

[7] H. T. Banks, *Identification of nonlinear delay systems using spline methods*, in "Nonlinear Phenomena in Mathematical Sciences" (ed. V. Lakshmikantham), Academic Press, New York, NY, (1982), 47–55.

[8] H. T. Banks, "A Functional Analysis Framework for Modeling, Estimation and Control in Science and Engineering," Chapman and Hall/CRC Press, Boca Raton, FL, 2012.

[9] H. T. Banks, J. E. Banks and S. L. Joyner, *Estimation in time-delay modeling of insecticide-induced mortality*, J. Inverse and Ill-posed Problems, **17** (2009), 101–125.

[10] H. T. Banks and D. M. Bortz, *A parameter sensitivity methodology in the context of HIV delay equation models*, J. Math. Biol., **50** (2005), 607–625.

[11] H. T. Banks, D. M. Bortz and S. E. Holte, *Incorporation of variability into the mathematical modeling of viral delays in HIV infection dynamics*, Math. Biosciences, **183** (2003), 63–91.

[12] H. T. Banks and J. A. Burns, *An abstract framework for approximate solutions to optimal control problems governed by hereditary systems,* in "International Conference on Differential Equations" (ed. H. Antosiewicz), Academic Press, New York, (1975), 10–25.

[13] H. T. Banks and J. A. Burns, *Hereditary control problems: Numerical methods based on averaging approximations*, SIAM J. Control & Opt., **16** (1978), 169–208.

[14] H. T. Banks, J. A. Burns and E. M. Cliff, *Parameter estimation and identification for systems with delays*, SIAM J. Control and Optimization, **19** (1981), 791–828.

[15] H. T. Banks, M. Davidian, J. R. Samuels, Jr. and Karyn L. Sutton, *An inverse problem statistical methodology summary*, in "Mathematical and Statistical Estimation Approaches in Epidemiology" (eds. Gerardo Chowell, et al.), Springer, Netherlands, (2009), 249–302.

[16] H. T. Banks, S. Dediu and S. L. Ernstberger, *Sensitivity functions and their uses in inverse problems*, Journal of Inverse and Ill-Posed Problems, **15** (2007), 683–708.

[17] H. T. Banks, S. Dediu, S. L. Ernstberger and F. Kappel, *A new approach to optimal design problems,* Journal of Inverse and Ill-Posed Problems, **18** (2010), 25–83.

[18] H.T. Banks, S. Dediu and H. K. Nguyen, *Time delay systems with distribution dependent dynamics,* IFAC Annual Reviews in Control, **31** (2007), 17–26.

[19] H. T. Banks, S. Dediu and H. K. Nguyen, *Sensitivity of dynamical systems to parameters in a convex subset of a topological vector space,* Math. Biosci. and Engr., **4** (2007), 403–430.

[20] H. T. Banks, K. Holm and F. Kappel, *Comparison of optimal design methods in inverse problems*, Inverse Problems, **27** (2011) 075002, 31 pp.

[21] H. T. Banks and F. Kappel, *Spline approximations for functional differential equations*, J. Differential Equations, **34** (1979), 496–522.

[22] H. T. Banks and J. M. Mahaffy, *Global asymptotic stability of certain models for protein synthesis and repression*, Quart. Applied Math., **36** (1978), 209–221.

[23] H. T. Banks and J. M. Mahaffy, *Stability of cyclic gene models for systems involving repression,* J. Theoretical Biology, **74** (1978), 323–334.

[24] H. T. Banks and H. Nguyen, *Sensitivity of dynamical systems to Banach space parameters,* J. Math. Analysis and Applications, **323** (2006), 146–161.

[25] H. T. Banks, Keri L. Rehm, Karyn L. Sutton, Christine Davis, Lisa Hail, Alexis Kuerbis and Jon Morgenstern, *Dynamic modeling of behavior change in problem drinkers*, N.C. State University, Center for Research in Scientific Computation technical report CRSC-TR11-08, Raleigh, NC, August, 2011; to appear in Quarterly of Applied Mathematics.

[26] H. T. Banks, D. Robbins and K. L. Sutton, *Theoretical foundations for traditional and generalized sensitivity functions for nonlinear delay differential equations*, N.C. State University, CRSC-TR12-14, Raleigh, NC, July, 2012.

[27] R. Bellman and K. L. Cooke, "Differential-Difference Equations," Mathematics in Science and Engineering, Vol. **6**, Academic Press, New York, NY, 1963.

[28] D. Brewer, *The differentiability with respect to a parameter of the solution of a linear abstract Cauchy problem*, SIAM J. Math. Anal., **13** (1982), 607–620.

[29] J. A. Burns, E. M. Cliff and S. E. Doughty, *Sensitivity analysis and parameter estimation for a model of Chlamydia Trachomatis infection*, J. Inverse Ill-Posed Problems, **15** (2007), 19–32.

[30] S. N. Busenberg and K. L. Cooke, eds., "Differential Equations and Applications in Ecology, Epidemics, and Population Problems," Academic Press, New York, 1981.

[31] V. Capasso, E. Grosso and S. L. Paveri-Fontana, eds., "Mathematics in Biology and Medicine," Lecture Notes in Biomath., Vol. **57**, Springer-Verlag, Berlin, Heidelberg, New York, 1985.

[32] J. Caperon, *Time lag in population growth response of* Isochrysis Galbana *to a variable nitrate environment*, Ecology, **50** (1969), 188–192.

[33] A. Casal and A. Somolinos, *Forced oscillations for the sunflower equation, entrainment*, Nonlinear Analysis: Theory, Methods and Applications, **6** (1982), 397–414.

[34] S. Choi and N. Koo, *Oscillation theory for delay and neutral differential equations*, Trends in Mathematics, **2** (1999), 170–176.

[35] K. L. Cooke, *Functional differential equations: Some models and perturbation problems,* in "Differential Equations and Dynamical Systems" (eds. J. K. Hale and J. P. LaSalle), Academic Press, New York, (1967), 167–183.

[36] J. M. Cushing, "Integrodifferential Equations and Delay Models in Population Dynamics," Lec. Notes in Biomath., Vol. **20**, Springer-Verlag, Berlin-New York, 1977.

[37] C. Elton, "Voles, Mice, and Lemmings," Clarendon Press, London, 1942.

[38] P. L. Errington, *Predation and vertebrate populations,* Quarterly Review of Biology, **21** (1946), 144–177, 221–245.

[39] U. Forys and A. Marciniak-Czochra, *Delay logistic equation with diffusion,* Proc 8th Nat. Conf. Mathematics Applied to Biology and Medicine, Lajs, Warsaw, Poland, (2002), 37–42.

[40] U. Forys and A. Marciniak-Czochra, *Logistic equations in tumor growth modelling,* Int.J. Appl. Math. Comput. Sci., **13** (2003), 317–325.

[41] J. S. Gibson and L. G. Clark, *Sensitivity analysis for a class of evolution equations*, J. Mathematical Analysis and Applications, **58** (1977), 22–31.

[42] L. Glass and M. C. Mackey, *Pathological conditions resulting from instabilities in physiological control systems*, Ann. N. Y. Acad. Sci., **316** (1979), 214–235.

[43] B. C. Goodwin, "Temporal Organization in Cells: A Dynamic Theory of Cellular Control Processes," Academic Press, London, 1963.

[44] B. C. Goodwin, *Oscillatory behavior in enzymatic control processes,* Adv. Enzyme Reg., **3** (1965), 425–439.

[45] K. Gopalsamy, "Stability and Oscillations in Delay Differential Equations of Population Dynamics," Mathematics and its Applications, **74**, Kluwer Academic Publishers Group, Dordrecht, 1992.

[46] K. P. Hadeler, *Delay equations in biology,* in "Functional Differential Equations and Approximation of Fixed Points" (eds. H. O. Peitgen and H. O. Walther), Lect. Notes Math., **730**, Springer-Verlag, Berlin, (1979), 136–156.

[47] J. Hale, "Theory of Functional Differential Equations," Second edition, Applied Mathematical Sciences, Vol. 3, Springer-Verlag, New York-Heidelberg, 1977.

[48] J. Hale and S. M. V. Lunel, "Introduction to Functional Differential Equations," Applied Math Sciences, Vol. **99**, Springer-Verlag, New York, 1993.

[49] F. C. Hoppensteadt, ed., "Mathematical Aspects of Physiology," Lectures in Applied Math, Vol. 19, American Mathematical Society, Providence, 1981.

[50] G. E. Hutchinson, *Circular causal systems in ecology*, Annals of the NY Academy of Sciences, **50** (1948), 221–246.

[51] G. E. Hutchinson, "An Introduction to Population Ecology," Yale University, New Haven, 1978.

[52] F. Kappel, *An approximation scheme for delay equations*, in "Nonlinear Phenomena in Mathematical Sciences" (ed. V. Lakshmikantham), Academic Press, New York, NY, (1982), 585–595.

[53] F. Kappel, *Generalized sensitivity analysis in a delay system*, Proc. Appl. Math. Mech., **7** (2007), 1061001–1061002.

[54] F. Kappel and W. Schappacher, *Autonomous nonlinear functional differential equations and averaging approximations*, J. Nonlinear Analysis, **2** (1978), 391–422.

[55] Y. Kuang, "Delay Differential Equations With Applications in Population Dynamics," Mathematics in Science and Engineering, Vol. **191**, Academic Press, New York, NY, 1993.

[56] N. MacDonald, *Time lag in a model of a biochemical reaction sequence with end-product inhibition*, J. Theor. Biol., **67** (1977), 727–734.

[57] N. MacDonald, "Time Lags in Biological Models," Lecture Notes in Biomathematics, **27**, Springer, Berlin, 1978.

[58] N. Minorsky, *Self-excited oscillations in dynamical systems possessing retarded actions,* Journal of Applied Mechanics, **9** (1942), A65–A71.

[59] N. Minorsky, *On non-linear phenomenon of self-rolling*, Proceedings of the National Academy of Sciences, **31** (1945), 346–349.

[60] N. Minorsky, "Nonlinear Oscillations," D. Van Nostrand, Co., Inc., Princeton, N.J.-Toronto-London-New York, 1962.

[61] D. M. Pratt, *Analysis of population development in Daphnia at different temperatures*, Biology Bulletin, **22** (1943), 345–365.

[62] D. Robbins, "Sensitivity Functions for Delay Differential Equation Models," Ph. D. Dissertation, North Carolina State University, Raleigh, NC, 2011.

[63] R. Schuster and H. Schuster, *Reconstruction models for the Ehrlich Ascites Tumor of the mouse*, in "Mathematical Population Dynamics, Vol. 2" (eds. by O. Arino, D. Axelrod and M. Kimmel), Wuertz, Winnipeg, (1995), 335–348.

[64] F. R. Sharpe and A. J. Lotka, *Contribution to the analysis of malaria epidemiology IV: Incubation lag,* supplement to Amer. J. Hygiene, **3** (1923), 96–112.

[65] Alfredo Somolinos, *Periodic solutions of the sunflower equation,* Quarterly of Applied Mathematics, **35** (1978), 465–478.

[66] K. Thomaseth and C. Cobelli, *Generalized sensitivity functions in physiological system identification*, Annals of Biomedical Engineering, **27** (1999), 607–616.

[67] E. M. Wright, *A non-linear difference-differential equation*, J. Reine Angew. Math., **494** (1955), 66–87.

Received July 05, 2012; revised November 14, 2012.

E-mail address: htbanks@ncsu.edu
E-mail address: danielle.evette.robbins@gmail.com
E-mail address: kls6479@louisiana.edu