Sensitivity Functions and Their Uses in Inverse Problems

H.T. Banks[1], Sava Dediu[2] and Stacey L. Ernstberger[3]

Center for Research in Scientific Computation
North Carolina State University
Raleigh, NC 27695-8205

Dedicated to Academician M. M. Lavrentiev on the occasion of his 75th birthday

July 21, 2007

Abstract

In this note we present a critical review of the some of the positive features as well as some of the shortcomings of the generalized sensitivity functions (GSF) of Thomaseth-Cobelli in comparison to traditional sensitivity functions (TSF). We do this from a computational perspective of ordinary least squares estimation or inverse problems using two illustrative examples: the Verhulst-Pearl logistic growth model and a recently developed agricultural production network model. Because GSF provide information on the relevance of data measurements for the identification of certain parameters in a typical parameter estimation problems, we argue that they provide the basis for new tools for investigators in design of inverse problem studies.

Keywords: Inverse problems, sensitivity and generalized sensitivity functions, Fisher information matrix, logistic growth model, agricultural production networks.

[1]htbanks@ncsu.edu
[2]sdediu@ncsu.edu
[3]sllawler@ncsu.edu

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
21 JUL 2007		**00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Sensitivity Functions and Their Uses in Inverse Problems	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **North Carolina State University,Center for Research in Scientific Computation,Raleigh,NC,27695-8205**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In this note we present a critical review of the some of the positive features as well as some of the shortcomings of the generalized sensitivity functions (GSF) of Thomaseth-Cobelli in comparison to traditional sensitivity functions (TSF). We do this from a computational perspective of ordinary least squares estimation or inverse problems using two illustrative examples: the Verhulst-Pearl logistic growth model and a recently developed agricultural production network model. Because GSF provide information on the relevance of data measurements for the identification of certain parameters in a typical parameter estimation problems, we argue that they provide the basis for new tools for investigators in design of inverse problem studies.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**33**	

1 Introduction

The interdisciplinary character of scientific research today leads to many instances in science, engineering and finance where mathematical models based on systems where either algebraic or differential equations (or both in hybrid models) are often used. In recent years, because of an increasing need to better describe social, physical or biological reality, and because of a tremendous improvement in computing technology, the models employed by scientists have became more complex and more sophisticated, wherein the number of variables and parameters being utilized have increased considerably. This naturally has motivated new questions of determining the relative importance of parameters, the effect on the model output of variation in parameters, the uncertainty in the model results due to the uncertainty of parameters, etc. Answers are not obvious for large, complex models and the associated problems provide significant new challenges for inverse problem research scientists. While the traditional literature on inverse problems has provided a wide range of useful theories and results on topics from well-posedness and regularity techniques to computational methodologies (e.g., see [23] and the extensive references therein), we focus here on concepts related to sensitivity which we believe can aid in experimental design to obtain data. These will enhance practical aspects of parameter estimation as well as motivate new questions in basic research. We do this in the context of several examples in which we introduce and illustrate ideas.

Sensitivity analysis is an ensemble of techniques [24] that can provide some answers to questions for a given problem of interest, yielding a much better understanding of the underlying mathematical model with a resulting marked improvement in the results obtained using the models. Traditionally, sensitivity analysis referred to a procedure used in simulation studies (direct problems) where one needed to evaluate the effects of parameter variations on the time course of model outputs and to identify the parameters or the initial conditions to which the model is most/least sensitive. In recent years however, due to an increasing interest in incorporating uncertainty into models and in ascertaining the sensitivity of parameter estimates with respect to data measurements, the uses of sensitivity have broadened significantly [7, 26]. In one direction, sensitivity of systems with probability measures embedded in the dynamics (problems involving aggregate dynamics) have become important in applications in biology, electromagnetics, and hysteretic and polymeric materials (see [7, 8] and the references therein). On

the other hand, investigators' attention has also recently turned to the sensitivity of the solutions to inverse problems with respect to data, in a quest for optimal selection of data measurements in experimental design. In this paper we focus on this latter direction.

As part of model validation and verification, one typically needs to estimate model parameters from data measurements, and a related question of paramount interest is related to sampling; specifically, at which time points the measurements are most informative in the estimation of a given parameter. Due to the fact that in practice the components of the parameter estimates are often correlated, traditional sensitivity functions (TSF) used alone are not efficient in answering this question because TSF do not take into account how model output variations affect parameter estimates in inverse problems. In an effort to overcome this shortcoming, Thomaseth and Cobelli [26] recently introduced a new class of sensitivity functions, called *generalized sensitivity functions* (GSF), which provide information on the relevance of measurements of output variables of a system for the identification of specific parameters. For a given set of time observations, Thomaseth and Cobelli use theoretical information criteria (the Fisher information matrix) to establish a relationship between the monotonicity of the GSF curves with respect to the model parameters and the information content of these observations. In this paper we present discussions on how to use this information content tool along with TSF to improve the parameter estimates in inverse problems and therefore to further validate the utility of these new functions in such problems. It is, of course, intuitive that sampling more data points from the region indicated by the GSF to be the "most informative" with respect to a given parameter would result in more information about that parameter, and therefore provide more accurate estimates for it.

We present a critical review of the positive features as well as shortcomings of the GSF, from the perspective of parameter estimation problems via several examples. We first illustrate ideas with the classic and well-known logistic growth model of Verhulst-Pearl. This is followed by discussion involving a recently developed model which we shall refer to as an agricultural production network model. Our paper is organized as follows. In Section 2 we introduce the necessary theoretical framework for our analysis and we define the traditional (TSF) and the generalized (GSF) sensitivity functions. In Sections 3 and 4 we introduce the Verhulst-Pearl logistic growth population model and the agricultural production network model, respectively, and we discuss numerical simulations carried out to investigate the effectiveness of

the information content provided by the GSF to improve the accuracy of the parameter estimates from the perspective of ordinary least square problems with noisy data. Finally, in Section 5 we present our conclusions and remarks and indicate directions for future work.

2 Theoretical Framework

We consider a parameter estimation problem for the general nonlinear dynamical system

$$\begin{aligned} \dot{x}(t) &= g(t, x(t), \theta) \\ x(t_0) &= x_0, \end{aligned} \tag{1}$$

with discrete time observations

$$y_j = f(t_j, \theta) + \epsilon_j = Cx(t_j, \theta) + \epsilon_j, \quad j = 1, \ldots, n \tag{2}$$

where $x, g \in \mathbb{R}^N$, $f, \epsilon_j \in \mathbb{R}^M$ and $\theta \in \mathbb{R}^p$. The matrix C is an $M \times N$ matrix which gives the observation data in terms of the components of the state variable x.

We assume for our model that the observation errors ϵ_j are independently identically distributed (i.i.d.), with zero mean. For different observation coordinates f_i, $i = 1, \ldots, M$, we have different variances σ_i^2 associated with the coordinates of the errors ϵ_j, i.e.

$$\epsilon_j \sim \mathcal{N}_M(0, V),$$

where $V = diag(\sigma_1^2, \ldots, \sigma_M^2)$. We also make the standard statistical assumption that there exists a "true" value parameter θ_0 such that the data set $\{y_j\}$, which can be interpreted as a realization of the observation process $Y = \{Y_j\}$ at the discrete time points t_j, $j = 1, \ldots, n$, has the form in equation (2) with $\theta = \theta_0$.

We use an ordinary least squares (OLS) approach to estimate θ_0, and we seek to find a value $\hat{\theta}^n$ that minimizes the cost functional

$$J_n(\theta) = \sum_{j=1}^{n} \left(f(t_j, \theta) - y_j\right)^T V^{-1} \left(f(t_j, \theta) - y_j\right). \tag{3}$$

Since $\{y_j\}$ is a realization of the random variable set $\{Y_j\}$, the estimate $\hat{\theta}^n$ we obtain by minimizing the cost functional J_n is a realization of some

random variable $\hat{\Theta}^n$. Therefore, the accuracy of our parameter estimates $\hat{\theta}^n$ ultimately depends on the statistical properties of this random variable, and in order to qualitatively analyze the estimates, we use a standard error approach [16, 19, 20, 25].

From the asymptotic theory of statistical analysis, one finds that as $n \to \infty$, $\hat{\Theta}^n \sim \mathcal{N}_p(\theta_0, \Sigma_0)$ is a good approximation, where the covariance matrix Σ_0 is given by

$$\Sigma_0 = \left(\sum_{j=1}^{n} D_j^T(\theta_0) V^{-1} D_j(\theta_0) \right)^{-1}, \tag{4}$$

and $D_j(\theta)$ is the $M \times p$ Jacobian matrix of f with respect to θ at t_j, i.e.,

$$D_j(\theta) = \frac{\partial f(t_j, \theta)}{\partial \theta}.$$

These are quite clearly the TSF for observations or measurement outputs with respect to the parameters. The covariance matrix Σ_0 is used in formulating the standard errors for our estimates $\hat{\theta}^n$; these are given by

$$\mathrm{SE}_k \ = \sqrt{(\Sigma_0)_{kk}}, \qquad k = 1, 2, ..., p. \tag{5}$$

Because θ_0 in (4) is unknown, we replace it by $\hat{\theta}^n$ when calculating approximations for (5). Moreover, the variances $\sigma_1^2, \ldots, \sigma_M^2$ are also generally unknown, and in order to use (4)-(5) in practice, we also use the following approximation

$$V \approx \hat{V} = \frac{1}{n-p} \sum_{j=1}^{n} [f(t_j, \hat{\theta}^n) - y_j][f(t_j, \hat{\theta}^n) - y_j]^T. \tag{6}$$

For the case when the observation system is scalar, i.e., $M = 1$, the $M \times M$ matrix V reduces to a scalar variance σ_0^2, and the equation (5) reduces to the standard formula

$$\mathrm{SE}_k \ = \sqrt{\hat{\sigma}^2 (\chi^T \chi)^{-1}_{kk}}, \qquad k = 1, 2, ..., p, \tag{7}$$

with $\chi(\theta)$ an $n \times p$ sensitivity matrix for our model given by

$$\chi_{jk}(\theta) = \frac{\partial f(t_j, \theta)}{\partial \theta_k}. \tag{8}$$

For this scalar observation system, the approximation $\hat{\sigma}^2$ to σ_0^2 in (7) is usually [16] taken as

$$\sigma_0^2 \approx \hat{\sigma}^2 = \frac{1}{n-p} \sum_{j=1}^{n} |f(t_j, \hat{\theta}^n) - y_j|^2. \tag{9}$$

2.1 Traditional Sensitivity Functions

Traditional sensitivity functions (TSF) are classical sensitivity functions used in mathematical modeling to investigate variations in the output of a model resulting from variations in the parameters and the initial conditions.

In order to quantify the variation in the state variable $x(t)$ with respect to changes in the parameter θ and the initial condition $x(t_0)$, we are naturally led to consider *(traditional) sensitivity functions (TSF)* defined by the derivatives

$$s_{\theta_k}(t) = \frac{\partial x}{\partial \theta_k}(t), \qquad k = 1, \ldots, p, \tag{10}$$

and

$$r_{x_{0l}}(t) = \frac{\partial x}{\partial x_{0l}}(t), \qquad l = 1, \ldots, N, \tag{11}$$

where x_{0l} is the l-th component of the initial condition x_0. If the function g is sufficiently regular, the solution x is differentiable with respect to θ_k and x_{0l}, and therefore the sensitivity functions s_{θ_k} and $r_{x_{0l}}$ are well defined.

Often in practice, the model under investigation is simple enough to allow us to combine the sensitivity functions (10) and (11), as is the case with the logistic growth population example discussed below. However, when one deals with a more complex model, as with the agricultural production network example, it is often preferable to consider these sensitivity functions separately for clarity purposes.

Because they are defined by partial derivatives which have a *local* character, the sensitivity functions are also local in nature. Thus sensitivity and insensitivity ($s_{\theta_k} = \partial x / \partial \theta_k$ very close to zero) depend on the time interval, the state values x, and the values of θ for which they are considered. Thus for example in a certain time subinterval we might find s_{θ_k} small so that the state variable x is *insensitive* to the parameter θ_k on that particular interval. The same function s_{θ_k} can take large values on a different subinterval, indicating to us that the state variable x is *very sensitive* to the parameter θ_k on the latter interval. From the sensitivity analysis theory for dynamical

systems, one finds that $s = (s_{\theta_1}, \ldots, s_{\theta_p})$ is an $N \times p$ vector function that satisfies the ODE system

$$\begin{aligned} \dot{s}(t) &= g_x(t, x(t), \theta)s(t) + g_\theta(t, x(t), \theta), \\ s(t_0) &= 0_{N\times p}, \end{aligned} \tag{12}$$

so that the dependence of s on $(t, x(t))$ as well as θ is readily apparent. Here we denote by $g_x = \partial g/\partial x$ and by $g_\theta = \partial g/\partial\theta$ the derivatives of g with respect to x and θ, respectively.

In a similar manner, the sensitivity functions with respect to the components of the initial condition x_0 define an $N \times N$ vector function $r = (r_{x_{01}}, \ldots, r_{x_{0N}})$, which satisfies

$$\begin{aligned} \dot{r}(t) &= g_x(t, x(t), \theta)r(t), \\ r(t_0) &= I_{N\times N}. \end{aligned} \tag{13}$$

The equations (12) and (13) are used in conjunction with equation (1) to numerically compute the sensitivities s and r for general cases when the function g is sufficiently complicated to prohibit a closed form solution by direct integration.

Because the parameters may have different units and the state variables may have varying orders of magnitude, sometimes in practice it is more convenient to work with the normalized version of the TSF, referred to as *relative sensitivity functions* (RSF). However, since in this paper we are using the standard error approach to analyze the performance of the least squares algorithm in estimating the true parameter values, we will focus solely on the non-scaled sensitivities, i.e., TSF.

2.2 Generalized Sensitivity Functions

Generalized sensitivity functions were proposed by Thomaseth and Cobelli [26] as a new tool in identification studies to analyze the distribution of the information content (with respect to the model parameters) of the output variables of a system for a given set of observations.

For a scalar observation model with discrete time measurements (i.e., when $M = 1$ and C is a $1 \times N$ array in (2)), the generalized sensitivity functions (GSF) are defined as

$$\mathbf{gs}(t_l) = \sum_{i=1}^{l} \frac{1}{\sigma^2(t_i)}[F^{-1} \times \nabla_\theta f(t_i, \theta_0)] \bullet \nabla_\theta f(t_i, \theta_0), \tag{14}$$

where $\{t_l\}$, $l = 1, \ldots, n$ are the times when the measurements are taken, and

$$F = \sum_{j=1}^{n} \frac{1}{\sigma^2(t_j)} \nabla_\theta f(t_j, \theta_0) \nabla_\theta f(t_j, \theta_0)^T \tag{15}$$

is the corresponding *Fisher information matrix*. The symbol "$\bullet$" represents element-by-element vector multiplication and, for motivation and details which led to the definition above, the interested reader may consult [11, 26]. The Fisher information matrix measures the information content of the data corresponding to the model parameters. In (14) we see that this information is transferred to the GSF, making them appropriate tools to indicate the relevance of the measurements in parameter estimation problems.

We note that the generalized sensitivity functions (14) are vector-valued functions with the same dimension as θ. The k-th component gs_k of the vector function **gs** represents the generalized sensitivity function with respect to θ_k. The GSF in (14) are defined only at the discrete time points $\{t_j, j = 1, \ldots, n\}$ and they are cumulative functions involving at time t_l only the contributions of those measurements up to and including t_l; thus gs_k calculates the influence of measurements up to t_l on the parameter estimate for θ_k.

It is readily seen from the definition that all the components of **gs** are one at the final time point t_n, i.e., $\mathbf{gs}(t_n) = \mathbf{1}$. If one defines $\mathbf{gs}(t) = \mathbf{0}$ for $t < t_1$ (naturally, **gs** is zero when no measurements are collected), then each component of **gs** varies between 0 and 1. As developed in [26], the time subinterval during which the change in gs_k has the *sharpest increase* corresponds to the observations which provide the most information in the estimation of θ_k. That is, regions of sharp increases in gs_k indicate a high concentration of information in the data about θ_k.

The numerical implementation of the generalized sensitivity functions (14) is straightforward, since the gradient of f with respect to θ (or x_0) is simply the Jacobian of x with respect to θ (or x_0) multiplied by the observation operator C. These Jacobian matrices can be obtained by numerically solving the sensitivity ODE system (12) or (13) coupled with the system (1). We will use this approach to compute the GSF for the agricultural production model. For the other example, the logistic model, the right side of equation (1) is sufficiently simple to permit a closed form solution.

3 The Verhulst-Pearl Population Model

We begin by considering the Verhulst-Pearl logistic growth equation [22], which approximates the evolution of a population size over time and is given by

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{K}\right). \tag{16}$$

The constants K and r represent the carrying capacity and the intrinsic growth rate respectively. The solution $x(t)$ of (16), representing the population number at time t, is given by

$$x(t) = \frac{K}{1 + \left(\frac{K}{x_0} - 1\right)e^{-rt}}, \tag{17}$$

where $x_0 = x(0)$ is the initial population size. The solution $x(t)$ approaches an asymptote at $x = K$ as $t \to \infty$; this is depicted in Figure 1.

The Verhulst-Pearl logistic equation is a relatively simple example with easily studied dynamics that is useful in demonstrating the utility of the traditional sensitivity functions as well as the generalized sensitivity functions in inverse problems (see [9] for more discussions on TSF for this system). Unless data is sampled from regions with changing dynamics, it is possible that some of the parameters will be difficult to estimate. Moreover, the parameters that are obtainable may have high standard errors as a result of introducing redundancy in the sampling region. In order to demonstrate this for the logistic growth problem, we will examine varying behavior in the model depending on the region from which t_j is sampled. We consider points τ_1 and τ_2, as depicted in Figure 1, partitioning the curve into three distinct regions: $0 < t_j < \tau_1$, $\tau_1 < t_j < \tau_2$, and $\tau_2 < t_j < T$, with T sufficiently large for our solution to be near its asymptote $x = K$. Based on the changing dynamics of the curve in Figure 1, we expect differences in the ability to estimate parameters depending on the region in which the solution is observed.

To illustrate these ideas, we carried out estimation procedures for the parameters $\theta = (K, r, x_0)$ in the logistic growth population model using ordinary least square procedures with both numerically generated (no-noise) and noisy simulated "data". We produced data sets $\{y_j\}_{j=1}^n$ by evaluating the numerical solution of (16) with the "true" value parameters θ_0 at t_j and then adding random noise in some cases. With a small dimension parameter

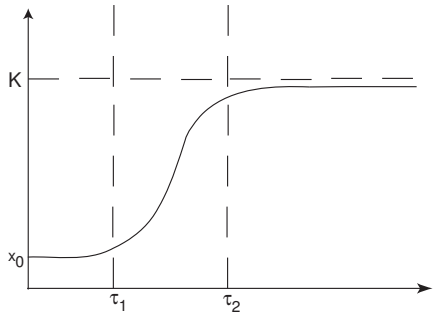

Figure 1: Distinct regions of growth in the Verhulst-Pearl solution curve.

space as in this example, a Nelder-Mead optimization algorithm is sufficient for the inverse problem. We hence use the MATLAB function $fminsearch$ to minimize the cost functional

$$J_n(\theta) = \sum_{j=1}^{n} |f(t_j, \theta) - y_j|^2 \tag{18}$$

with respect to θ and using an initial guess θ^0 for the optimization algorithm. In order to avoid what is typically called an *inverse crime*, we evaluate f in the cost function using the ODE solver $ode15s$ with (16), which returns the numerical approximation to the solution $f(t, \theta) = x(t; K, r, x_0)$, rather than using the analytical solution (17). This, in effect, produces "noisy data" even when no additional noise is added. In this note we illustrate ideas with a specific example, taking $\theta_0 = (K, r, x_0) = (17.5, 0.7, 0.1)$; the corresponding solution curve $x(t)$ can be seen in Figure 2(a).

3.1 Traditional Sensitivities

We consider an ordinary least squares problem for the estimation of the parameters $\theta = (K, r, x_0)$ in the logistic growth model (16), using the explicit solution given by (17) and then examining the sensitivities with respect to

the parameters K, r, and x_0. We can readily compute the partial derivatives

$$\begin{aligned} \frac{\partial x}{\partial K} &= \frac{x_0^2(1-e^{-rt})}{(x_0+(K-x_0)e^{-rt})^2}, \\ \frac{\partial x}{\partial r} &= \frac{Kx_o(K-x_0)te^{-rt}}{(x_0+(K-x_0)e^{-rt})^2}, \\ \frac{\partial x}{\partial x_0} &= \frac{K^2e^{-rt}}{(x_0+(K-x_0)e^{-rt})^2}, \end{aligned} \tag{19}$$

which are the traditional sensitivity functions s_K, s_r and s_{x_0}.

We analyze the TSF corresponding to each parameter in the initial region of the curve, where the solution approaches x_0. When we consider the initial region of the curve, where $0 < t_j < \tau_1$ for $j = 1, \ldots, n$, we have

$$\frac{\partial x(t_j)}{\partial K} \approx 0, \qquad \frac{\partial x(t_j)}{\partial r} \approx 0, \qquad \frac{\partial x(t_j)}{\partial x_0} \approx 1;$$

this follows from considering the limits of the sensitivity functions (19) as $t \to 0$. Based on the above analytical findings, which indicate low sensitivities with respect to K and r, we expect to have little ability to determine these parameters when we sample data from $[0, \tau_1]$; however we should be able to estimate x_0.

We next consider the region of the curve which is near the asymptote at $x = K$, in this case for $\tau_2 < t_j < T$, $j = 1, \ldots, n$. Here we find that by considering the limits as $t \to \infty$, we have the approximations

$$\frac{\partial x(t_j)}{\partial K} \approx 1, \qquad \frac{\partial x(t_j)}{\partial r} \approx 0, \qquad \frac{\partial x(t_j)}{\partial x_0} \approx 0.$$

Based on these approximations, we expect to be able to estimate K well when we sample data from $[\tau_2, T]$. However, using data only from this region, we do not expect to be able to estimate x_0 or r.

Finally, we consider the part of the solution curve where $\tau_1 < t_j < \tau_2$ for $j = 1, \ldots, n$ and where it has nontrivially changing dynamics. We note that the partial derivative values differ greatly from the values in regions $[0, \tau_1]$ and $[\tau_2, T]$. When $[\tau_1, \tau_2]$ is included in the sampling region we expect to recover good estimates for all three parameters.

Our analytical observations are fully consistent with information contained in the graphs of the TSF illustrated in Figure 2(b) for $T = 25$. We

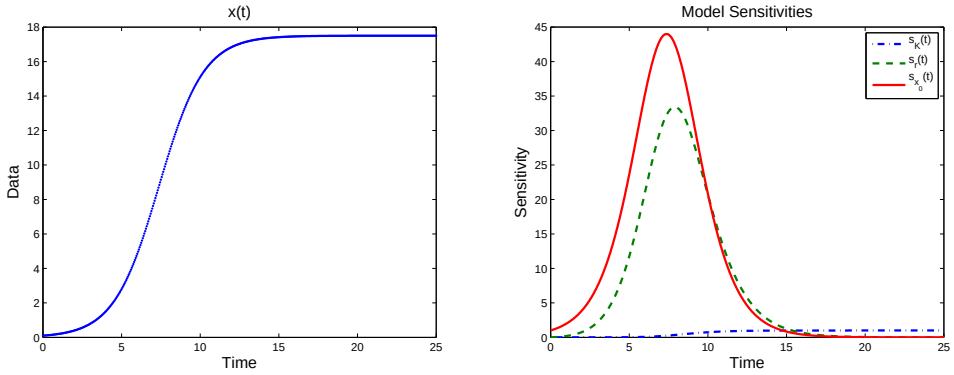

Figure 2: (a) Logistic curve (b) TSF corresponding to each parameter for the logistic curve with $\theta = (17.5, 0.7, 0.1)$.

note that the curve s_K slowly increases with time and it appears that the solution is insensitive to K until around the flex point of the logistic curve, which occurs shortly after $t = 7$ in this case. The sensitivities s_K and s_r both are close to zero when t is near the origin, and hence we deduce that both K and r will be difficult or impossible to obtain using data in that region. Also, we observe that s_{x_0} and s_r are nearly zero in [15,25], which suggests that we will be unable to estimate x_0 or r using observations in that region.

In order to computationally illustrate how the traditional sensitivity theory applies to our logistic growth example, we present here a summary of findings obtained using numerically (no-noise added) simulated data generated with $ode15s$. We also restricted the time domain from which we sampled data points to one of the three regions of interest $[0, \tau_1]$, $[\tau_1, \tau_2]$ or $[\tau_2, T]$ for each inverse problem calculation. Then we determined the standard errors of each parameter set in order to analyze the success of the algorithm at approximating the various components of θ. (Other computations with varying levels of added noise in the "data" are presented in [9] where findings are similar to those reported here.)

We first sampled data from the region where the solution curve is close to x_0, and for this example we considered the region $[0, 1]$. We used several initial guesses in our MATLAB solver, and expected with each guess that we would be able to achieve an appropriate estimate for x_0, but perhaps not for r or K. In actuality we were able to obtain close estimates for both r and

x_0, as reported in Table 1.

$\boldsymbol{\theta^0}$	$\hat{\boldsymbol{\theta}}$	**Standard Errors**
Values obtained using data in $0 \leq t \leq 1$		
(8,1,0.3)	(10.686,0.7038,0.09998)	(10.398,0.003589,3.4206e-05)
(15,2,0.2)	(26.8241,0.6978,0.1000)	(7.4758,0.002581,2.4593e-05)
(30,0.3,0.5)	(34.1993,0.6970,0.1000)	(9.9947,0.003450,3.2879e-05)
(10,0.1,0.3)	(16.9140,0.7001,0.1000)	(0.7924,0.000274,2.6068e-06)
Values obtained using data in $1 \leq t \leq 15$		
(8,1,0.3)	(17.4985,0.6998,0.1001)	(0.002062,0.0003049,0.0002142)
(15,2,0.2)	(17.4985,0.6998,0.1001)	(0.002063,0.0003050,0.0002143)
(30,0.3,0.5)	(17.4985,0.6998,0.1001)	(0.002061,0.0003048,0.0002141)
(10,0.1,0.3)	(17.4984,0.6998,0.1001)	(0.002063,0.0003051,0.0002144)
Values obtained using data in $15 \leq t \leq 25$		
(8,1,0.3)	(17.4879,1.1427,0.0495)	(0.02912,1.5016,2.3000)
(15,2,0.2)	(17.4877,1.7838,0.1908)	(0.03076,1.5865,2.4301)
(30,0.3,0.5)	(17.5017,0.5887,0.5374)	(0.00268,0.1381,0.2115)
(10,0.1,0.3)	(17.5023,0.5468,1.0098)	(0.00328,0.1689,0.2587)

Table 1: The optimized θ values and corresponding standard errors, on the given intervals with $\theta_0 = (17.5, 0.7, 0.1)$ implemented using a computationally noisy data set.

Upon further examination we found that r and x_0 are highly correlated when $0 \leq t \leq 1$, which is evident by the magnitude of their correlation coefficients, given in Table 2. By studying each iteration of the MATLAB solver, we observed that a good estimate for x_0 was easily obtained, and then, due to the high correlation between r and x_0, r was eventually obtained each time. However, true to our predictions, K was never reasonably estimated using data from this region.

Recall that correlation coefficients range from -1 to 1, where higher magnitudes indicate stronger correlation between the corresponding parameters. The correlation coefficients displayed in Table 2 were computed using the standard definition

$$corr(X, Y) = \frac{cov(X, Y)}{\sigma_X \sigma_Y}$$

where X and Y are two random variables with mean μ_X and μ_Y and variances

	K	r	x_0
K	1.0000	-0.3434	0.2855
r	-0.3434	1.0000	-0.977
x_0	0.2855	-0.977	1.0000

Table 2: Correlation coefficients for parameters

σ_X^2 and σ_Y^2. The parameter estimates $\hat{\theta}^n$ can be interpreted as realizations of a random variable $\hat{\Theta}^n$, with distribution which, from the asymptotic theory of statistical analysis for the least square algorithm, can be approximated by a normal distribution as $n \to \infty$, i.e., for n large, $\hat{\Theta}^n \sim \mathcal{N}_p(\theta_0, \sigma_0^2(\chi^T\chi)^{-1})$ is a good approximation. The correlation coefficients between two components of θ are thus given by

$$corr(\theta_k, \theta_l) = \frac{cov(\theta_k, \theta_l)}{\sigma_{\theta_k}\sigma_{\theta_l}}.$$

Using the definition of the covariance matrix, we have that $cov(\theta_k, \theta_l)$ is simply the (k,l)-th element of $\sigma_0^2(\chi^T\chi)^{-1}$ and the standard deviations σ_{θ_k} and σ_{θ_l} are the square roots of the (k,k)-th and (l,l)-th diagonal entries. This was used to compute the approximate correlation coefficients of Table 2.

The logistic model above was considered and analyzed in [9] using traditional sensitivity functions. However the model was formulated with a different parameterization: instead of $\theta = (K, r, x_0)$, the parameter set was given by $\beta = (a, b, x_0) = (r, \frac{r}{K}, x_0)$. Studying the TSF curves with this other parameterization, we found there was no difficulty in predicting the regions in which the state was sensitive to each parameter with no consideration given to the correlation coefficients. With a new parameterization, $\theta = (K, r, x_0)$, the correlation between the coefficients reveals direct information on our ability to predict the identifiability for each parameter. Comparing these findings with those of [9], we see that *the parameterization of the model is critical in sensitivity studies.*

We next considered the region where the dynamics of the curve are changing, and for our example this region is the interval $[1, 15]$. As expected, regardless of the initial guess that is used in the solver, we can generally obtain reasonable estimates for K, r and x_0; this can be seen in Table 1.

Finally we sampled data from the region where the curve approaches an asymptote at K, and for this example we considered the region $[15, 25]$. We

used several initial guesses in our MATLAB solver, and expected that with each guess we would be able to achieve an appropriate estimate for K, but not for r or x_0. This is precisely what occurred with every initial guess.

We see that the TSF, used in conjunction with the correlation coefficients, provided sufficient information in these examples to predict which parameters could be determined using data from different regions. Similar results obtained using noise-added data with the logistic model, along with some pitfalls of the TSF, can be found in [9].

3.2 Generalized Sensitivities

We next illustrate the utility of the generalized sensitivity functions by applying the theory to the logistic growth model (16). We start by numerically computing the GSF using equation (14) with $\sigma = 1$ and the true value parameters $\theta_0 = (17.5, 0.7, 0.1)$. The plots of these functions are shown in Figure 3(b) where one can observe obvious regions of steep increase in each curve. For the curves $gs_{x_0}(t)$, $gs_r(t)$ and $gs_K(t)$, we find by visual inspection

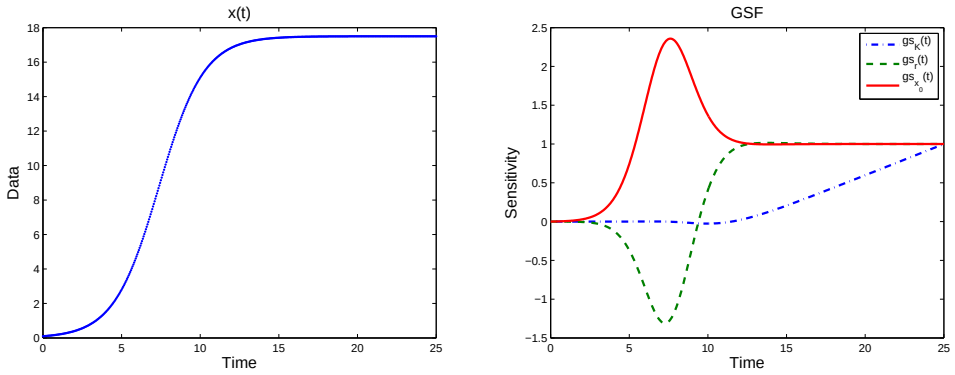

Figure 3: (a) Logistic curve (b) GSF corresponding to each parameter for the logistic curve with $\theta = (17.5, 0.7, 0.1)$.

that these regions are approximately $[4.5, 7.5]$, $[7, 11]$ and $[12, 25]$, respectively. By the aforementioned generalized sensitivity theory, if we increase the number of data points sampled in one of these regions, the estimation of the corresponding parameter is expected to improve.

In order to implement the GSF for the logistic growth example, we introduced noise into the simulated data by adding Gaussian noise $\epsilon_j \sim \mathcal{N}(0, \sigma_0^2)$

for $j = 1, \ldots, n$. To obtain the results in Tables 3, 4, and 5, we used a randomly generated noisy data set with $\sigma_0 = 0.5$. We initially sampled n

n	m	t_{GSF}	t_{nonGSF}
50	0	(0.10748,0.029371,0.021604)	(0.10748,0.029371,0.021604)
50	25	(0.077068,0.027991,0.020752)	(0.10441,0.021464,0.015622)
50	50	(0.059962,0.025752,0.019184)	(0.099866,0.01754,0.012632)
125	0	(0.061814,0.016807,0.01236)	(0.061814,0.016807,0.01236)
125	25	(0.053069,0.016569,0.012234)	(0.061051,0.014428,0.010561)
125	50	(0.046431,0.016095,0.011919)	(0.06453,0.013726,0.010003)
125	125	(0.037958,0.016276,0.012124)	(0.061772,0.010767,0.0077615)
250	0	(0.047175,0.012805,0.0094158)	(0.047175,0.012805,0.0094158)
250	50	(0.039627,0.012356,0.009123)	(0.046264,0.010884,0.0079648)
250	100	(0.035866,0.012422,0.009198)	(0.045259,0.0095868,0.0069842)
250	250	(0.027388,0.011737,0.0087426)	(0.04239,0.0073602,0.0053052)

n	m	$t_{uniform}$
50	0	(0.10748,0.029371,0.021604)
50	25	(0.086565,0.023623,0.017374)
50	50	(0.077352,0.021049,0.01548)
125	0	(0.061814,0.016807,0.01236)
125	25	(0.057056,0.015504,0.011401)
125	50	(0.053388,0.014502,0.010664)
125	125	(0.047175,0.012805,0.0094158)
250	0	(0.047175,0.012805,0.0094158)
250	50	(0.043685,0.011854,0.0087166)
250	100	(0.0404,0.01096,0.0080592)
250	250	(0.034008,0.0092229,0.0067818)

Table 3: This table shows standard errors corresponding to the addition of m extra points in or excluding [12,25], as it is the period of steepest increase for the parameter K.

data points uniformly over the entire region $[0, 25]$, and then we sampled an additional m points from varying parts of the entire region, comparing the standard errors of each trial.

Each parameter in θ has a specific interval of steepest increase according to the corresponding GSF, and the m additional points are sampled according to those regions. In each of Tables 3, 4,and 5, we consider three separate time grids:

- t_{GSF} consists of n uniform time points over [0,25] with m points added in the area of steepest increase according to the GSF for each parameter,

- t_{nonGSF} consists of n uniform time points over [0,25] with m points added everywhere except the area of steepest increase according to the GSF for each parameter,

- $t_{uniform}$ is the time grid referring to n uniform time points over [0,25] with m additional points added as uniformly as possible over the entire region.

For example, in Table 3, the t_{GSF} column refers to the standard errors generated by optimizing θ over n uniformly distributed data points in [0,25] with m additional data points sampled from the steepest increase interval corresponding to K: [12,25]. The standard errors in the t_{nonGSF} column refer to the same n initial data points, with the additional m points sampled from outside the GSF suggested region for K.

In general, values of standard errors are meaningful only relative to the estimated values of the corresponding parameters. Thus, one should report both the estimated parameter values and the corresponding SE. However, here and in Section 4, we report only the changes in SE as data sets are changed. This is because our primary focus is on these changes in SE and because the corresponding OLS estimates are near (same order magnitude) the true values θ_0; hence it suffices to simply report only θ_0 and the changes in SE.

Note that since $\theta = (K, r, x_0)$, the standard error for K is indicated as the first entry in each of the ordered sets in each table, i.e., $SE_K = SE_{\theta_1}$, and Table 3 refers to the steepest region of increase in the GSF corresponding to K. Similarly Tables 4 and 5 use the GSF regions corresponding to r and x_0, respectively, and so SE_{θ_2} and then SE_{θ_3} are the entries of interest.

We would have expected that for each parameter, the corresponding standard error in the t_{GSF} column would generally improve as m increased. In actuality, the corresponding standard errors generally improved when m additional points were sampled in all three grids for each parameter. However, we get better results with the t_{GSF} grid than with the $t_{uniform}$ grid, and the t_{nonGSF} grid was the worst of the three. Hence we see that by catering to the GSF-recommended regions, we are able to obtain better standard errors for

n	m	t_{GSF}	t_{nonGSF}
50	0	(0.10748,0.029371,0.021604)	(0.10748,0.029371,0.021604)
50	25	(0.097735,0.021034,0.016379)	(0.085987,0.025298,0.01814)
50	50	(0.10215,0.019931,0.015575)	(0.071516,0.022395,0.015627)
125	0	(0.061814,0.016807,0.01236)	(0.061814,0.016807,0.01236)
125	25	(0.061485,0.014585,0.011165)	(0.055069,0.0155,0.011258)
125	50	(0.063405,0.013979,0.010842)	(0.052744,0.015359,0.011012)
125	125	(0.062507,0.012136,0.009478)	(0.043153,0.013462,0.0093937)
250	0	(0.047175,0.012805,0.0094158)	(0.047175,0.012805,0.0094158)
250	50	(0.046673,0.011057,0.0084657)	(0.043127,0.012158,0.0088167)
250	100	(0.045837,0.010087,0.0078221)	(0.038498,0.01117,0.0080133)
250	250	(0.043884,0.0085094,0.0066447)	(0.032496,0.010137,0.0070706)

n	m	$t_{uniform}$
50	0	(0.107480,0.029371,0.021604)
50	25	(0.086565,0.023623,0.017374)
50	50	(0.077352,0.021049,0.015480)
125	0	(0.061814,0.016807,0.012360)
125	25	(0.057056,0.015504,0.011401)
125	50	(0.053388,0.014502,0.010664)
125	125	(0.047175,0.012805,0.009416)
250	0	(0.047175,0.012805,0.0094158)
250	50	(0.043685,0.011854,0.0087166)
250	100	(0.040400,0.010960,0.0080592)
250	250	(0.034008,0.009223,0.0067818)

Table 4: This table shows standard errors corresponding to the addition of m extra points in or excluding [7,11], as it is the period of steepest increase for the parameter r.

the corresponding parameters than if we had merely increased the number of points sampled over the entire region.

3.3 "Forced-to-one" Artifact

We first note that the shape of the TSF curves remains the same regardless of the amount of data that is sampled, whereas the GSF curves are data dependent and change shape with varying amounts of data. We can see in Figure 4 that when we restrict the data set for the logistic growth model to

n	m	t_{GSF}	t_{nonGSF}
50	0	(0.10748,0.029371,0.021604)	(0.10748,0.029371,0.021604)
50	25	(0.10338,0.021906,0.014617)	(0.081923,0.02323,0.017621)
50	50	(0.099638,0.018577,0.012139)	(0.07291,0.021106,0.016235)
125	0	(0.061814,0.016807,0.01236)	(0.061814,0.016807,0.01236)
125	25	(0.064238,0.015175,0.010452)	(0.056836,0.015457,0.011486)
125	50	(0.059416,0.012951,0.0086999)	(0.052824,0.01476,0.011091)
125	125	(0.05867,0.010902,0.0071197)	(0.043777,0.012942,0.0099389)
250	0	(0.047175,0.012805,0.0094158)	(0.047175,0.012805,0.0094158)
250	50	(0.046931,0.011043,0.0076218)	(0.043069,0.011892,0.0088565)
250	100	(0.046695,0.010136,0.006819)	(0.040509,0.011467,0.0086329)
250	250	(0.043801,0.0081056,0.0052964)	(0.031893,0.0095317,0.0073385)

n	m	$t_{uniform}$
50	0	(0.10748,0.029371,0.021604)
50	25	(0.086565,0.023623,0.017374)
50	50	(0.077352,0.021049,0.01548)
125	0	(0.061814,0.016807,0.01236)
125	25	(0.057056,0.015504,0.011401)
125	50	(0.053388,0.014502,0.010664)
125	125	(0.047175,0.012805,0.0094158)
250	0	(0.047175,0.012805,0.0094158)
250	50	(0.043685,0.011854,0.0087166)
250	100	(0.04040,0.010960,0.0080592)
250	250	(0.034008,0.0092229,0.0067818)

Table 5: This table shows standard errors corresponding to the addition of m extra points in or excluding [4.5,7.5], as it is the period of steepest increase for the parameter x_0.

$[0, 2]$, the TSF curves look the same as those where we merely zoom in to $[0, 2]$ after sampling data over the entire region $[0, 25]$. Then in Figure 5 the GSF curves look very different when the sampled data is only from the interval $[0, 2]$ as compared to when the data is sampled from $[0, 25]$ and zoomed-in to $[0, 2]$. If an insufficient amount of data is used for parameter estimations, the GSF curves can be misleading because by definition the GSF are forced to be equal to one by the end of the data set. This so-called "forced-to-one" artifact can cause misleading regions of steep increase in the GSF curves as can be seen in the curve gs_K when $t \in [1.7, 2]$ in Figure 5(a). Observe that in

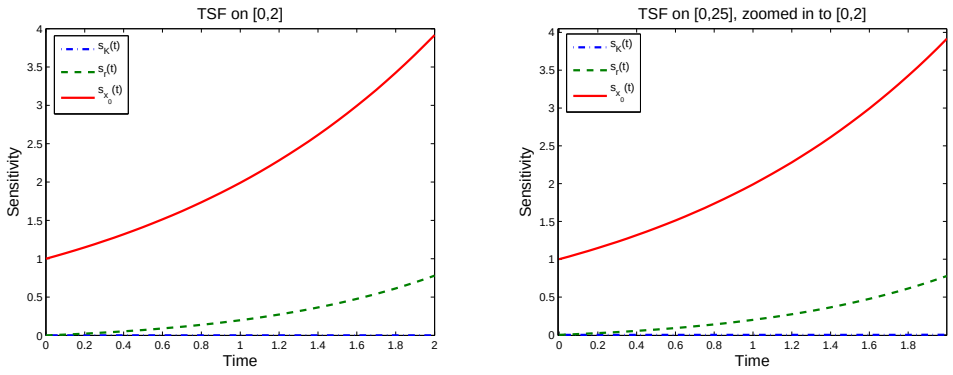

Figure 4: In (a) and (b) we see the TSF of the parameters when sampling data in [0,2] and the zoomed-in portion when data was observed from the entire region [0,25].

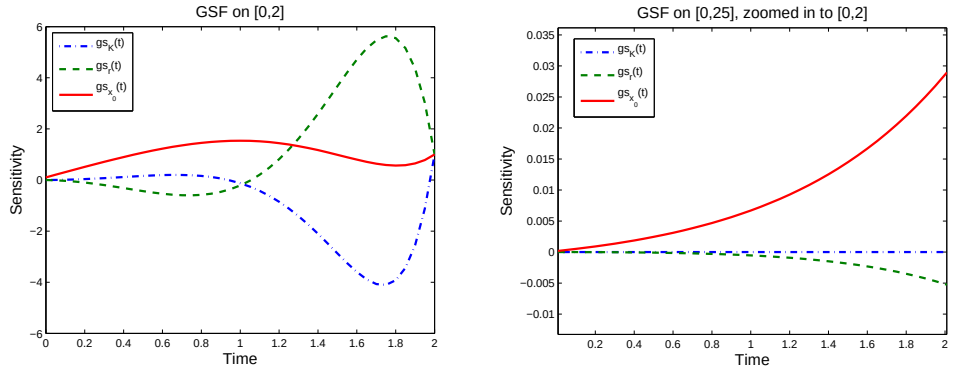

Figure 5: In (a) we see the GSF of the parameters when sampling data in [0,2] and in (b) we have the zoomed-in portion [0,2] when data was observed from the entire region [0,25].

Figure 4 we can see that the state is clearly not sensitive to the K parameter in region $[0, 2]$, and hence sampling additional data points in the period of misleading increase would merely make estimates worse, with an increase in standard error. We can also see another example of this in Figure 6 where we compare the GSF curves generated from data sampled from $[0, 0.2]$ to the curves generated when data is sampled from $[0, 25]$. In Figure 6(b) it is clear

that there is no significant period of increase for either K or r, however in (a) it appears that the corresponding GSF curves both have distinct regions of (again in this case misleading) increase. Therefore, it is important to note that while the GSF curves can be misleading when a limited portion of data is obtainable, the TSF curves can still be used appropriately in such sensitivity studies.

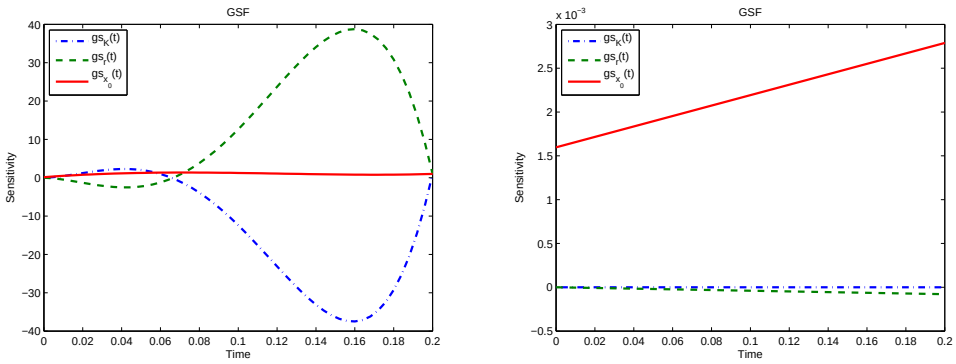

Figure 6: In (a) we see the GSF of the parameters when sampling data in [0,0.2] and in (b) we have the zoomed-in portion [0,0.2] when data was observed from the entire region [0,25].

It is also important to remark on the Fisher information matrix which appears in the definition of generalized sensitivity functions (14). The information on the parameters provided by the measurement y_i is quantified by the derivative of an information index with respect to θ_i. When we try to estimate the parameters r and x_0 with data from the interval $[15, 25]$ alone (see Table 1), we obtain large errors which increase in magnitude when the number of sample points increases. Although initially unexpected, this phenomenon can be explained based on the GSF discussion presented above. When we sample additional data points from the region where the traditional sensitivity curves are flat and the generalized sensitivity functions exhibit the "forced-to-one" artifact, we actually introduce redundancy in the sensitivity matrix. This considerably increases the condition number of the Fisher information matrix, which in turn, by the Crammer-Rao inequality, causes the variance of the unbiased estimator to be very large, making our estimates less useful.

Although there is improvement when the GSF-recommended regions are considered, the amount of additional points sampled to garner the improved standard errors needs to be taken into consideration. Depending on the problem, the cost may be too high to sample additional data points for the slightly improved results. However, in other cases the additional sampling may be worth the improvement. While a useful tool, the GSF may not be an efficient choice in some parameter improvement attempts.

4 An Agricultural Production Model

We continue our investigations on the relevance of the generalized sensitivity functions to parameter estimations problems, in this case with a more complex example of an agricultural production network model formulated in terms of the nonlinear dynamical system

$$\begin{aligned}
\dot{c}_1(t) &= -\kappa_1 c_1(t)(l_2 - c_2(t))_+ + \kappa_4 \min(c_4(t), s_m) \\
\dot{c}_2(t) &= -\kappa_2 c_2(t)(l_3 - c_3(t))_+ + \kappa_1 c_1(t)(l_2 - c_2(t))_+ \\
\dot{c}_3(t) &= -\kappa_3 c_3(t)(l_4 - c_4(t))_+ + \kappa_2 c_2(t)(l_3 - c_3(t))_+ \\
\dot{c}_4(t) &= -\kappa_4 \min(c_4(t), s_m) + \kappa_3 c_3(t)(l_4 - c_4(t))_+
\end{aligned} \tag{20}$$

together with the initial conditions

$$\mathbf{c}(0) = \mathbf{c}_0. \tag{21}$$

The system (20) is the deterministic limit for large populations (in a sense made precise in [2]) of a continuous time discrete state Markov Chain proposed in [2] to model the flow and the impact of eventual disturbances in a swine production network. For simplicity, the modeled network is assumed to consist of four levels of production nodes: Growers (N_1), Nurseries (N_2), Finishers (N_3), and Processing Plants (N_4), and each node represents an aggregation of all the production units corresponding to that level in the production process. Although unrealistic, it is assumed that there are no losses in the first three nodes of the production network and the only deaths occur at the processing plants. Another important assumption is that the network is *closed*, i.e., there is a direct flow from node N_4 to node N_1, instantly replenishing the network and keeping the total population size constant in time. We note that this assumption is realistic when the network is efficient

and operates at or near full capacity (i.e., when the number of animals removed from the chain are immediately replaced by new production/growth, avoiding significant idle times).

The state variables $c_i(t)$, $i = 1, \ldots, 4$, represent the swine population size (roughly speaking, an ensemble average in the sense of the Markov Chain model of [2]) at the nodes N_i at time t. The parameters κ_i, $i = 1, \ldots, 4$, in (20) represent the transition rates between consecutive nodes and l_i, $i = 2, 3, 4$, represent the maximum capacitates at the nodes N_i. There is no capacity constraint at node N_1, but there is a maximum slaughtering capacity at node N_4 (processing plants) which we denote by s_m. For any real z, the symbol $(z)_+$ is defined as the positive part of z, i.e., $(z)_+ = \max(z, 0)$. The numerical values of all the parameters which we used in our analysis presented here are listed in Table 6. For more details about the model (20) and its derivation, the interested reader may consult [2].

Parameters	Definition	Values	Units
κ_1	scaled rate at node 1	2.879	1/days
κ_2	scaled rate at node 2	1.093	1/days
κ_3	scaled rate at node 3	1.51	1/days
κ_4	scaled rate at node 4	1	1/days
l_2	scaled capacity at node 2	$2.387 \cdot 10^{-1}$	dimensionless
l_3	scaled capacity at node 3	$6.498 \cdot 10^{-1}$	dimensionless
l_4	scaled capacity at node 4	$5.67 \cdot 10^{-2}$	dimensionless
s_m	scaled slaughter capacity	$1.039 \cdot 10^{-1}$	dimensionless
$c_1(0)$	scaled initial condition	$9.45 \cdot 10^{-2}$	dimensionless
$c_2(0)$	scaled initial condition	$2.221 \cdot 10^{-1}$	dimensionless
$c_3(0)$	scaled initial condition	$6.313 \cdot 10^{-1}$	dimensionless
$c_4(0)$	scaled initial condition	$5.19 \cdot 10^{-2}$	dimensionless

Table 6: Aggregated agricultural network model: Parameters for deterministic simulations (numbers of pigs are in thousands).

We considered a series of least square problems using simulated noisy data in order to probe the utility of the GSF pertaining to estimation problems. Since we have three categories of parameters in our model (transition rates, capacities and initial conditions of the network) we tried to estimate the parameters in each category when all the others remain fixed. The simulated

data was generated by first numerically solving the system (20) with parameter and initial condition values given in Table 6 (we will refer to these values as the *true parameter values* in these discussions), and then adding Gaussian noise of zero mean and standard deviation 0.1 to the solution obtained at each observation point.

We begin with the problem of estimating the transition rates κ_1, κ_2, κ_3 and κ_4 when the l and c_{0k} parameters are fixed at the values given in Table 6. For the true parameter values $\theta_0 = \bar{\kappa} = (2.88, 1.09, 1.51, 1)$, we plot the generalized sensitivity functions (14) and the traditional sensitivity functions in Figure 7. In both of these plots, one can observe a distinct time

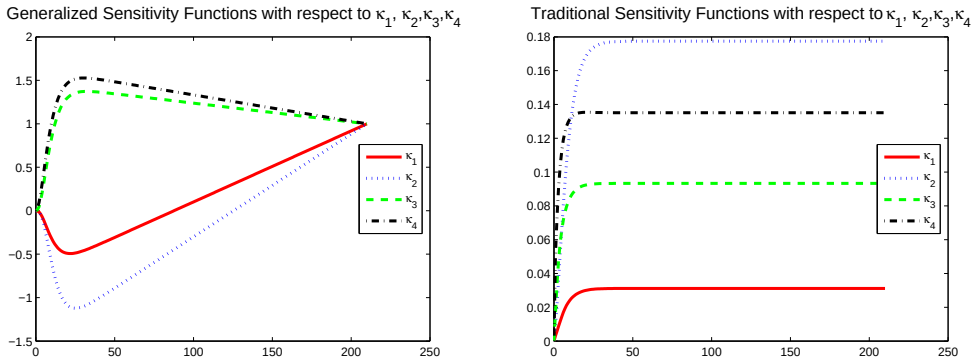

Figure 7: Generalized and Traditional Sensitivity Functions with respect to κ_1, κ_2, κ_3, κ_4 corresponding to parameters values given in Table 6.

point near $t = 30$, where the dynamics of both the GSF and TSF curves change. Between 0 and this point, the TSF plots exhibit a sharp monotonic increase, and then they reach a steady state very quickly. On the contrary, the GSF curves increase/decrease steeply before reaching this time point, and then are forced to one. According to the GSF theory, the approximate interval $[0, 30]$ is the region in which measurements are the most informative for estimating the true parameters $\bar{\kappa}$. This means that by sampling additional data points here, we expect to obtain more information about $\bar{\kappa}$, resulting in more accurate estimates for these parameters.

By comparing the GSF plots, we also observe that strong correlations exist between κ_1, κ_2, κ_3 and κ_4. The correlation coefficients between these parameters, given in Table 7, reflect the dynamics of the curves shown in

	κ_1	κ_2	κ_3	κ_4
κ_1	1.000	0.922	0.879	0.902
κ_2	0.922	1.000	0.933	0.910
κ_3	0.879	0.933	1.000	0.730
κ_4	0.902	0.910	0.730	1.000

Table 7: Correlation coefficients for each parameter

	Data points in			Standard Errors for			
Data Set	[0,30]	[40,210]	Total	κ_1	κ_2	κ_3	κ_4
KDS_1	10	35	45	0.315	0.108	0.166	0.101
KDS_2	30	35	65	0.205	0.067	0.114	0.062
KDS_3	10	57	67	0.244	0.089	0.136	0.085
KDS_4	10	43	53	0.296	0.094	0.140	0.094
KDS_5	10	86	96	0.288	0.103	0.148	0.096
KDS_6	10	172	182	0.254	0.093	0.131	0.087
KDS_7	10	0	10	0.551	0.177	0.313	0.166
KDS_8	30	0	30	0.240	0.075	0.140	0.066
KDS_9	0	35	35	3659.1	1343.9	1860.3	1195.1
KDS_{10}	0	86	86	2364.5	871.9	1189.3	800.47
True Value Parameters				**2.88**	**1.09**	**1.51**	**1**

Table 8: Typical standard errors for the transition rates κ_1, κ_2, κ_3 and κ_4 with a series of different types of data sets.

Figure 7, and also support our intuitive reasoning about flows in closed networks functioning at capacity. For such networks, the transition rates from one state to the other are obviously highly correlated.

In order to illustrate the above theory, we again used ordinary least squares inverse problems with numerous sets of data. We performed the least square minimization for several data sets of type KDS_1 with a total of 45 observations, where 10 are uniformly distributed within the interval $[0, 30]$ and 35 are uniformly distributed within the interval $[40, 210]$ (for simplicity we exclude the transition interval $[30, 40]$ from our present analysis). Standard errors for the estimates of κ_1, κ_2, κ_3 and κ_4 in a typical optimization

with one of the KDS_1 data sets are shown in Table 8.

When increasing the number of points sampled in the interval $[0, 30]$ to 30 and keeping the number of data points in $[40, 210]$ the same (the KDS_2 type data set, Table 8), the standard errors corresponding to each of the parameters κ_1, κ_2, κ_3 and κ_4 decrease considerably. The new standard errors range between 61% and 68% of the standard errors for the data set KDS_1. A decrease in the standard errors is also observed when we solve the least square problem using data (sets of type KDS_7, KDS_8 of Table 8) only from the interval $[0, 30]$, where the standard errors obtained with data set KDS_8 range between 40% and 45% of the standard errors for KDS_7. Thus, numerical calculations fully support the GSF theory that increasing the number of data points in the region $[0, 30]$ results in more accurate estimates for the parameter $\bar{\kappa}$.

Next we increased the number of data points sampled from the interval $[40, 210]$, and obtained an entirely different outcome. By comparing the entries for the data sets KDS_4, KDS_5 and KDS_6 in Table 8, we see that there is little improvement in the standard errors when we successively solve the least square problem with a fixed number of 10 data points in the interval $[0, 30]$ and 43, 86 and respectively 172 data points in $[40, 210]$. Also, when we attempt to estimate the parameters using *only* data (data sets KDS_9 and KDS_{10}) sampled from the region $[40, 210]$, the resulting standard errors are large (the parameter estimates were also quite bad for all of the KDS_9 and KDS_{10} type data sets), increasing in magnitude as the number of sampled points increases. As in the case of the logistic growth model above, this is not surprising if one recalls the investigations in [9]. It is an expected consequence of introducing redundancy in the sensitivity matrix, which in turn makes the Fisher information matrix ill-posed, yielding huge standard errors.

We also performed a similar analysis for the estimation of the nodal capacities l_i and of the initial conditions c_{0k} from data with Gaussian noise added as described above. In Figures 8 and 9 we depict the generalized and traditional sensitivity functions with respect to l_2, l_3, l_4 and c_{01}, c_{02}, c_{03} and c_{04}, respectively. In both cases, one can distinguish a small time interval, at the beginning of the time axis, where the traditional sensitivity functions exhibit sharp increases/decreases before reaching the steady state. The same time interval also gives the region where the generalized sensitivity functions with respect to c_{0k} exhibit sharp increases/decreases before following the "forced-to-one" artifact. Based on the theory, we can anticipate that for the

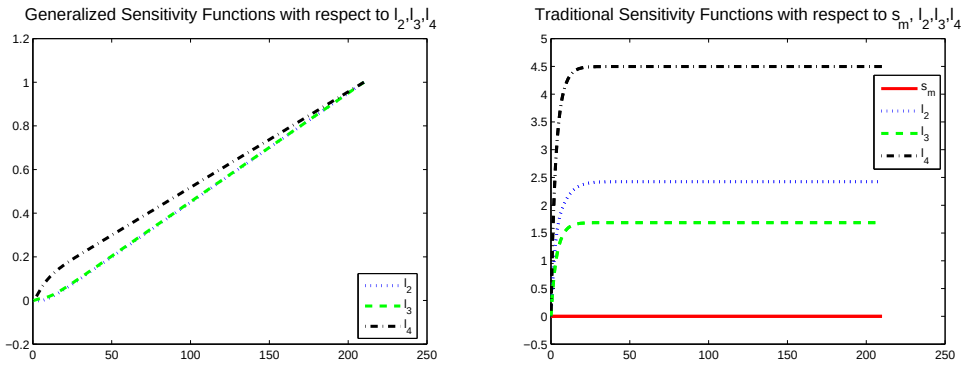

Figure 8: Generalized and Traditional Sensitivity Functions for s_m, l_2, l_3, l_4.

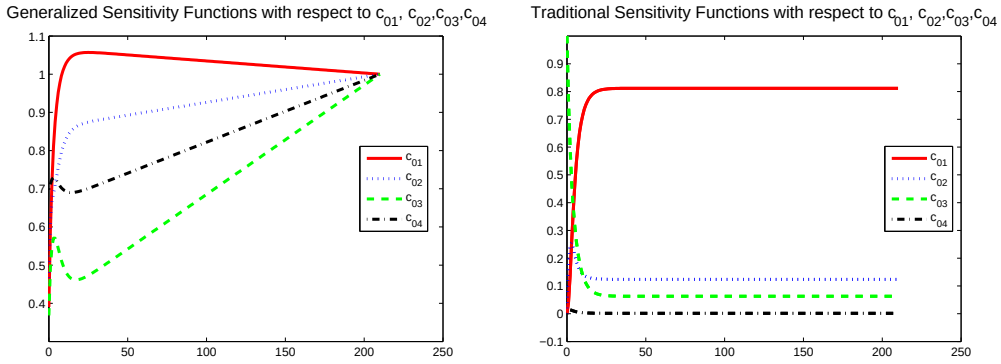

Figure 9: Generalized and Traditional Sensitivity Functions for c_{01}, c_{02}, c_{03}, c_{04}.

estimation of the initial conditions c_{0k}, the data sampled from this initial time interval is the most informative. Numerical calculations presented in Table 9 confirm this expectation. Indeed, increasing the number of data points in the interval $[0, 24]$ leads to smaller residual errors and better estimates (see the standard errors corresponding to CDS_2 and CDS_8), whereas increasing the number of data points in the interval $[34, 210]$ does not provide much improvement (see CDS_4, CDS_5 and CDS_6). Similar to the findings when we estimated the κ's, when we use only data points from the interval $[34, 210]$ for the estimation of c_{0k} (see CDS_9 and CDS_{10}), we obtain very large stan-

	Data points in			Standard Errors for			
Data Set	[0,24]	[34,210]	Total	c_{01}	c_{02}	c_{03}	c_{04}
CDS_1	8	36	44	0.0075	0.0079	0.0077	0.0094
CDS_2	24	36	60	0.0059	0.0064	0.0062	0.0088
CDS_3	8	60	68	0.0083	0.0089	0.0079	0.0099
CDS_4	8	45	53	0.0087	0.0084	0.0074	0.0097
CDS_5	8	89	97	0.0089	0.0086	0.0077	0.0095
CDS_6	8	178	186	0.0082	0.0088	0.0076	0.0092
CDS_7	8	0	8	0.0110	0.0116	0.0062	0.0124
CDS_8	24	0	24	0.0062	0.0071	0.0065	0.0090
CDS_9	0	36	36	1596.50	1434.50	879.97	1544.90
CDS_{10}	0	89	89	630.24	859.35	483.58	1095.80
True Values for c_{0k}				**0.0945**	**0.2221**	**0.6313**	**0.0519**

Table 9: Typical standard errors for the initial conditions c_{0k} with a series of different types of data sets.

dard errors (as well as unacceptable parameter estimates) compared to the previous ones.

As established in [2], s_m is the parameter to which the system (20) is the least sensitive overall. In fact for the particular values used for simulations ($s_m = 0.1039$, $l_2 = 0.2387$, $l_3 = 0.6498$, $l_4 = 0.0567$) the output of the system (20) does not depend on s_m at all (see the traditional sensitivity functions from Figure 8). From the particular form of our system and by practical insight, this can be explained intuitively by the fact that when the capacity s_m of the first node is too high, the flow c_4 (which replenishes the network) will never reach it, making this capacity constraint inactive for our dynamical system. The consequence is that for the given data, the entries corresponding to s_m in the sensitivity matrix are all zero, making the Fisher information matrix singular. Intuitively, this simply says that it is impossible to reconstruct a parameter from data where the solution values *do not depend at all* on that parameter. This is the reason for which in Figure 8 we present the generalized sensitivity functions only with respect to l_2, l_3 and l_4. Unlike the generalized sensitivity functions for κ and c_{0k} which present an initial region with sharp increases/decreases followed by a region where they exhibit a "forced-to-one" artifact, the GSF for l_2, l_3 and l_4 show a

	Data points in			Standard Errors for		
Data Set	[0,24]	[44,210]	Total	l_2	l_3	l_4
LDS_1	8	34	42	0.0023	0.0062	0.0031
LDS_2	24	34	58	0.0019	0.0051	0.0026
LDS_3	8	56	64	0.0018	0.0049	0.0025
LDS_4	8	42	50	0.0020	0.0056	0.0028
LDS_5	8	84	92	0.0014	0.0041	0.0020
LDS_6	8	168	176	0.0011	0.0030	0.0015
LDS_7	8	0	8	0.0071	0.0165	0.0086
LDS_8	24	0	24	0.0032	0.0077	0.0039
LDS_9	0	34	34	0.0026	0.0073	0.0037
LDS_{10}	0	84	84	0.0015	0.0044	0.0021
True Values for l_2, l_3 and l_4				**0.2387**	**0.6498**	**0.0567**

Table 10: Typical standard errors for the capacities l_2, l_3 and l_4 with a series of different types of data sets.

steady increase from zero to one throughout their time courses. We anticipate therefore that for the estimation of the capacities l, the information is more uniformly distributed in the interval $[0, 210]$. Numerical results presented in Table 10, where higher number of data points result in lower standard errors, confirm these expectations.

We conclude this section with the observation that the analysis and the numerical simulations reported here illustrate the utility of the GSF in investigations of reasonably complex estimation problems. However, as in the case of the logistic growth model, we emphasize that one should use care when making inferences from the regions of monotonicity for these functions. In addition to the regions of genuine information content, the GSF often indicate regions of steep increase (the interval $[40, 210]$ in Figure 7 and the interval $[34, 210]$ in Figure 9) which are due simply to the "forced-to-one" artifact inherent in the GSF as defined in [26].

5 Conclusions

From the analysis presented in this paper and in [11, 26], it is clear that generalized sensitivity functions are useful tools in inverse problem investigations. The primary positive feature of generalized sensitivity functions is their ability to suggest regions of high information content where, if additional observations are taken, one can generally expect to improve specific parameter estimates. This is supported by our numerical findings for the logistic growth model and a recently developed agricultural production network model, and further supports the initial computational findings with examples in [26]. Moreover, by visually investigating the dynamics of TSF and GSF curves, we can potentially identify subsets of parameters which are highly correlated. Although the GSF theory alone can be of use in formulation of parameter estimation problems, the most insight can be gained when the GSF are used in conjunction with traditional sensitivity functions. This will ensure maximum understanding of parameter sensitivity and can be of great value in improving data acquisition design.

Acknowledgments

This research was supported in part (HTB) by the National Institute of Allergy and Infectious Disease under grant 9R01AI071915-05, in part (HTB and SD) by the U.S. Air Force Office of Scientific Research under grant AFOSR-FA9550-04-1-0220, in part (SD) by the Statistical and Applied Mathematical Sciences Institute, which is funded by NSF under grant DMS-0112069, and in part (SLE) by the U.S. Dept of Homeland Security with a Fellowship in the DHS Scholarship and Fellowship Program, a program administered by the Oak Ridge Institute for Science and Education (ORISE) for DHS through an interagency agreement with the U.S Department of Energy (DOE). ORISE is managed by Oak Ridge Associated Universities under DOE contract number DE-AC05-06OR23100. All opinions expressed in this paper are the authors' and do not necessarily reflect the policies and views of DHS, DOE, or ORAU/ORISE.

References

[1] H. M. Adelman and R.T. Haftka, Sensitivity analysis of discrete structural systems, *A.I.A.A. Journal*, **24** (1986), 823–832.

[2] P. Bai, H. T. Banks, S. Dediu, A. Y. Govan, M. Last, A. L. Loyd, H. K. Nguyen, M. S. Olufsen, G. Rempala, and B. D. Slenning, Stochastic and deterministic models for agricultural production networks, CRSC-TR07-06, NCSU, February, 2007; *Math. Biosci. and Engr.*, **4** (2007), 373–402.

[3] H.T. Banks and D. M. Bortz, A parameter sensitivity methodology in the context of HIV delay equation models, CRSC-TR02-24, NCSU, August, 2002; *J. Math. Biology*, **50** (2005), 607–625.

[4] H.T. Banks and D. M. Bortz, Inverse problems for a class of measure dependent dynamical systems, CRSC-TR04-33, NCSU, September, 2004; *J. Inverse and Ill-posed Problems*, **13** (2005), 103–121.

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

[6] H.T. Banks, D.M. Bortz, G.A. Pinter and L.K. Potter, Modeling and imaging techniques with potential for application in bioterrorism, Chapter 6 in *Bioterrorism: Mathematical Modeling Applications in Homeland Security*, (H.T. Banks and C. Castillo-Chavez, eds.), Frontiers in Applied Mathematics **FR28**, SIAM, Philadelphia, 2003, pp. 129–154.

[7] H. T. Banks, S. Dediu and H.K. Nguyen, Sensitivity of dynamical systems to parameters in a convex subset of a topological vector space, CRSC-TR06-25, NCSU, September, 2006; *Math. Biosci. and Engineering,* **4** (2007), 403–430.

[8] H. T. Banks, S. Dediu and H. K. Nguyen, Time delay systems with distribution dependent dynamics, CRSC-TR06-15, May, 2006; *IFAC Annual Reviews in Control*, **31** (2007), 17–26.

[9] H. T. Banks, S. L. Ernstberger and S. L.Grove, Standard errors and confidence intervals in inverse problems: sensitivity and associated pitfalls,

CRSC-TR06-10, NCSU, March, 2006; *J. Inverse and Ill-posed Problems*, **15** (2007), 1–18.

[10] H. T. Banks and H. K. Nguyen, Sensitivity of dynamical system to Banach space parameters, CRSC Tech Rep. CRSC-TR05-13, NCSU, February, 2005; *J. Math. Anal. Appl.*, **323** (2006), 146–161.

[11] J. J. Batzel, F. Kappel, D. Schneditz and H.T. Tran, *Cardiovascular and Respiratory Systems: Modeling, Analysis and Control*, SIAM Frontiers in Applied Math, SIAM, Philadelphia, 2006.

[12] L.M.A. Bettencourt, A. Cintron-Arias, D.I. Kaiser and C. Castillo-Chavez, The power of a good idea: quantitative modeling of the spread of ideas from epidemiological models, Preprint No. LAUR-05-0485, Los Alamos National Laboratory, January, 2005.

[13] S.M. Blower and H. Dowlatabadi, Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model as an example, *International Statistics Review*, **62** (1994), 229–243.

[14] G. Casella and R. L. Berger, *Statistical Inference,* Duxbury, California, 2002.

[15] J.B. Cruz, ed., *System Sensitivity Analysis*, Dowden, Hutchinson & Ross, Inc., Stroudsberg, PA, 1973.

[16] M. Davidian and D. Giltinan, *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall, London, 1998.

[17] M. Eslami, *Theory of Sensitivity in Dynamic Systems: An Introduction*, Springer-Verlag, Berlin, 1994.

[18] P.M. Frank, *Introduction to System Sensitivity Theory*, Academic Press, Inc., New York, NY, 1978.

[19] A. R. Gallant, *Nonlinear Statistical Models*, John Wiley & Sons, Inc., New York, 1987.

[20] R. I. Jennrich, Asymptotic properties of non-linear least squares estimators., *Ann. Math. Statist.*, **40** (1969), 633–643.

[21] M. Kleiber, H. Antunez, T.D. Hien and P. Kowalczyk, *Parameter Sensitivity in Nonlinear Mechanics: Theory and Finite Element Computations*, John Wiley & Sons, New York, NY, 1997.

[22] M. Kot, *Elements of Mathematical Ecology*, Cambridge University Press, Cambridge, 2001.

[23] M. M. Lavrentiev and L. Ya. Saveliev, *Operator Theory and Ill-Posed Problems*, Koninklijke Brill NV, Leiden, The Netherlands, 2006.

[24] A. Saltelli, K. Chan and E.M. Scott, eds., *Sensitivity Analysis*, Wiley Series in Probability and Statistics, John Wiley & Sons, New York, NY, 2000.

[25] G. A. F. Seber and C. J. Wild, *Nonlinear Regression,* John Wiley & Sons, Inc., New York, 1989.

[26] K. Thomaseth, C. Cobelli. Generalized sensitivity functions in physiological system identification., *Ann. Biomed. Eng.*, **27** (1999), 607–616.