ESTIMATING RATE EQUATIONS USING NONPARAMETRIC REGRESSION METHODS

Kristen Meier

Institute of Statistics Mimeograph Series No. 1969

December, 1989

TABLE OF CONTENTS

1. INTRODUCTION

Estimation of derivatives for noisy data is important in identifying and distinguishing biological processes. We consider observed data of the form (t_i, y_i) that is assumed to follow the model

$$y_i = f(t_i) + e_i \qquad i=1,\ldots,n$$

where f is a regression(smooth) function, and the e_i are independent and identically distributed, mean zero errors. Given such data, there are several possible ways to estimate $f'(t)$. These methods are discussed in section 3.3. However, in many biological applications, what is of direct interest to the biologist is not $f'(t)$, but the *relationship* between f(t) and $f'(t)$. That is, biologists are interested in estimating the function Φ where $\Phi(f) = -f'$. There might or might not exist an appropriate parametric model, say $\Phi_{\underset{\sim}{\theta}}$, for the function Φ. We would like to estimate f' in such a way that we can also obtain a good estimate of Φ.

One application that has motivated this research is the estimation of rate equations using data that has been collected from a *solution depletion experiment*. This type of experiment is often used to quantify and study a plants ability to absorb different ions. A detailed description of such an experiment is given in section 2. An example is provided from the study of a corn seedlings ability to absorb nitrogen (Jackson, 1989). Specifically, the experimenter was interested in the relationship between the nitrate concentration of a solution surrounding the roots of a corn seedling and the seedling roots rate of uptake (absorption) of nitrate. (Nitrogen is an essential mineral for green plants (Keeton, 1976)). However, in a solution depletion experiment, the uptake rate is not measured directly. Instead, the ion concentration of the surrounding solution is measured over time, the rate of uptake being determined by minus the derivative of this curve.

If C(t) denotes the ion concentration of a solution surrounding plant roots at time t, then $-\frac{d}{dt}C(t)$ is the uptake rate *as a function of time t*. To estimate Φ, where $\Phi(C) = -C'$, we need the uptake rate as a function of concentration. At concentration $u=C(t)$, the rate of uptake *as a function of concentration u* is $-\frac{d}{dt}C(t)$ evaluated at $t=C^{-1}(u)$. Therefore, we define the *uptake curve* or *rate equation* of interest to be

$$-\Phi(u)=\frac{d}{dt}C(t)\Big|_{t=C^{-1}(u)}=C'(C^{-1}(u))$$

where $\Phi(u)$ is the uptake rate of the ion when the ion concentration of the surrounding solution is $u=C(t)$.

Knowledge of the functional form of Φ is important in quantifying the ability of a particular type of plant to absorb a specific ion. In addition, scientists would like to estimate certain features of a rate curve, e.g. the maximum rate of uptake, and to compare features of two rate curves, e.g. the features of two different varieties or the features of one variety that has been subjected to two different treatments.

My research will address the problem of how to estimate Φ and to examine the properties of such an estimate. I will show that by using nonparametric regression methods, in particular by using smoothing splines (defined in section 3), one can obtain an actual curve estimate (versus a point estimate) of the rate curve Φ and this estimate is consistent (section 4.1). In addition, I suggest that meaningful features of the rate curve can be estimated using the rate curve estimate and that such features of different rate curves can then be compared. These features may or may not correspond to parameters in a proposed parametric form for Φ. In the case that there does exist a reasonable parametric model for Φ, say $\Phi_{\underset{\sim}{\theta}}$, I conjecture that my proposed estimate of Φ will yield weakly consistent estimates of the parameters $\underset{\sim}{\theta}$. I further conjecture that the conventional method, which uses first differences to estimate the derivative, yields estimates of $\underset{\sim}{\theta}$ that are not consistent. These conjectures are based on some preliminary results discussed in section 4.2.

2. SOLUTION DEPLETION EXPERIMENTS

A solution depletion experiment is carried out in the following (simplified) manner: A cup is filled with a solution that has a predetermined ion concentration, c_0. At time zero, the roots of a plant are placed in the solution. Then, at each of n predetermined time points, $t_1, t_2, \ldots t_n$, the solution surrounding the roots is sampled and the concentration, $y_1, y_2, \ldots y_n$ is recorded. A simplified view is that over time, the external concentration of the ion decreases; this decrease is due to the *uptake* or absorption of the ion by the plant roots.

It is assumed that the data follow the model

$$y_i = C(t_i) + e_i \qquad i=1,\ldots,n$$

where C, the *depletion curve*, is a smooth, monotonically decreasing function and the e_i's are random measurement error. An example of such data is given in Figure 1a. Theoretically, the rate of uptake is determined by taking minus the time derivative of the depletion curve C. Since C is unkown, the experimenters estimated this derivative by taking the difference in consecutive concentrations divided by the time elapsed. That is, they computed the rate as $r_i = -(y_i - y_{i-1})/(t_i - t_{i-1})$. However, this method gave very noisy estimates of rates and the experimenters were concerned about using this transformed data, (y_i, r_i) $i=2,\ldots n$, to further estimate the *uptake curve* Φ, where $\Phi(C) = -C'$. The results of this method are shown in Figure 1b. In section 3.3 we discuss more sophisticated numerical methods to estimate a derivative from noisy data.

Running this same experiment under several different conditions, a scientist can gather a great deal of information about a particular plant. For example, running this experiment using different initial concentrations, c_0, gives insight into whether or not a plant has "memory". That is, it may give insight into the question: does a plants rate of uptake at concentration c^* depend on the concentration to which the plant was initially exposed at time zero? In addition, this experiment can be run using plants that have been subjected to different pretreatments. One pretreatment might be different exposure levels to different ions. Finally, the exact same experiment can be run using different varieties of a particular plant to give insight into how the varieties differ.

3. NONPARAMTRIC REGRESSION ESTIMATES

Recently, methods have been developed where a regression curve can be estimated without imposing a parametric model. Instead one assumes only that the regression curve is "smooth," that is that it satisfies certain differentiability conditions. Such methods are called ***nonparametric regression methods*** because the models are "parameter-free." These methods are not to be confused with ***classical nonparametric methods*** which are often labeled "distribution-free" methods and typically involve the order statistics. In the remainder of this paper, the word nonparametric will imply parameter-free models.

3.1 Defining the basic model

The basic model considered in nonparametric regression is

$$y_i = f(t_i) + e_i \qquad i=1,\dots,n$$

where $\underset{\sim}{y}$ is an observed vector depending on a smooth function f, on the vector of points $\underset{\sim}{t}$ and on independent and identically distributed, mean zero errors $\underset{\sim}{e}$ with variance equal to σ^2. For simplicity, we also assume that $0=t_1 \le t_2 \le \dots \le t_n = 1$ and that $f:[0,1] \rightarrow \mathbf{R}$. The goal is to estimate f by assuming only that it is smooth; specification of a parametric form for f is not required.

One method of estimation is called ***kernel estimation***. Priestley and Chao (1972) proposed the first kernel estimators for f in the case where $\underset{\sim}{t}$ is fixed. Nadaryra (1964) and Watson (1964) developed an estimator for the random design case. The idea is to estimate f(x) by a weighted average of observations near x. Weights are specified by a *kernel function*, K, and "nearness to x" is specified by a *bandwidth* or *smoothing parameter*, b ($=b_n$). We consider the ***kernel estimate*** of the regression function f at x:

$$\hat{f}_b(x)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b}K\left(\frac{x-t_i}{b}\right)y_i$$

where K is a kth order kernel if it satisfies the following moment conditions

$$\int_{[-1,1]} K(u)\,du = 1,$$

$$\int_{[-1,1]} u^j K(u)\,du = 0, \qquad j=1,\ldots,m-1,$$

$$\int_{[-1,1]} u^m K(u)\,du = \alpha \neq 0,$$

$$\int_{[-1,1]} [K(u)]^2\,du < \infty.$$

The bandwidth, b controls the relative weight given to an observation based on its distance from x. Large values of b yield smooth estimates of f by averaging over a wider neighborhood of points while small values of b yield more "wiggly" estimates of f. For a general discussion of kernel estimators, refer to Müller (1988) and Eubank (1988).

My research will concentrate on another method of estimation called *smoothing spline estimation.* Stimulated by the works of Whittaker (1923), Schoenberg (1964) derived the smoothing spline estimator. Later Reinsch (1967) gave an independent derivation for a special case (m=2) of the estimator. The idea is to find an estimator of f that fits the data well *and* that is smooth. Goodness-of-fit is typically measured by $n^{-1}\sum_{i=1}^{n}(y_i-f(t_i))^2$. If we assume that $f \in W_2^m[0,1]$, where $W_2^m[0,1]=\{f: f,\ldots f^{(m-1)}$ absolutely continuous, $f^{(m)} \in L_2[0,1]\}$, then a natural measure of smoothness associated with f is $\int_0^1 (f^{(m)}(u))^2 du$. Combining these two measures, an *m^{th}- order smoothing spline estimate* of the regression function f, denoted $\hat{f}_\lambda$, is the $f \in W_2^m[0,1]$ that minimizes

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - f(t_i))^2 + \lambda \int_{[0,1]} (f^{(m)}(u))^2 du.$$

The solution, $\hat{f}_\lambda$ is a piecewise polynomial of order $2m-1$. Typically $m=2$, in which case $\hat{f}_\lambda$ is a piecewise cubic polynomial. We denote $J_m(f)=\int_0^1 (f^{(m)}(u))^2 du$ as the *penalty term* and we call λ $(=\lambda_n)$ the ***smoothing parameter*** because it controls the "smoothness" of the solution versus the goodness-of-fit of the solution. A large value of λ places importance on smoothness and penalizes estimators with a large mth derivative; as $\lambda \to \infty$, the spline estimate approaches the linear least squares estimate. Conversely, a small value of λ places more emphasis on goodness-of-fit; as $\lambda \to 0$, the spline estimate approaches an interpolation spline.

Since $\hat{f}_\lambda$ is a piecewise polynomial of order 2m-1, it can be rewritten as a linear combination of *B-splines*, basis functions that span the space of piecewise polynomials of order 2m-1 (deBoor, 1972). This correspondence allows us to compute smoothing spline estimates using efficient algorithms developed for B-splines. Related to this fact, one can use the Riesz Representation Theorem in conjunction with the fact that W_2^m is a reproducing kernel Hilbert space to show that for a fixed value of λ there exists a matrix $A(\lambda)$ such that

$$\hat{f}_\lambda(\underset{\sim}{t}) = A(\lambda)\underset{\sim}{y}$$

where $\hat{f}_\lambda(\underset{\sim}{t})=[\hat{f}_\lambda(t_1)\ \hat{f}_\lambda(t_2)\ \ldots\ \hat{f}_\lambda(t_n)]^T$. That is, like a kernel estimator, the smoothing spline point estimator is a linear combination of the data points. This representation has motivated research in approximating the smoothing spline estimate by a kernel estimate. See Nychka (1989).

Studying the theoretical properties of kernel estimators is more tractable than studying the theoretical properties of smoothing splines. Therefore, most of my preliminary results have been shown using kernel estimates. Using the results of Nychka (1989), these results will be extended to smoothing splines.

3.2 Choosing the smoothness parameter

Both kernel methods and smoothing spline methods require some procedure to choose their smoothing parameters, b and λ, respectively. Choosing this parameter objectively is an important issue that has been and continues to be widely discussed in the literature. The ideas used in each case are similar, therefore I will discuss one of these methods used to obtain λ, the smoothing parameter for smoothing splines. For a detailed discussion on selecting a bandwidth for kernel estimates, see Härdle, Hall and Marron (1988). We assume that σ^2, the error variance, is unknown. Th reader is referred to Mallows(1973) and Craven and Wahba (1979) for the case when σ^2 is known.

One criterion that is used to obtain a value for λ is based on an unbiased estimate of the prediction error. A value of λ is called optimal if it minimizes the *expected average squared error* (Craven and Wahba, 1979), also called the *risk function* (Rice, 1984), where average squared error is defined to be

$$\mathrm{ASE}(\lambda)=\frac{1}{n}\sum_{i=1}^{n}\left(\hat{f}_\lambda(t_i)-f(t_i)\right)^2.$$

Taking the expected value and using the fact that $\hat{f}_\lambda(\underset{\sim}{t})=A(\lambda)\underset{\sim}{y}$, we have

$$E(\mathrm{ASE}(\lambda))=E\,\frac{1}{n}\,\|A(\lambda)\underset{\sim}{y}-f(\underset{\sim}{t})\|^2$$

where the norm is the Euclidean norm. The optimal λ, denoted $\hat{\lambda}_0$ is the minimizer of $E(ASE(\lambda))$. In practice, we can not determine $\hat{\lambda}_0$ if we do not know f. However, the method of ***cross-validation*** can be used to obtain an estimate of this optimal lambda.

The idea behind cross-validation is to choose a value of λ so that $\hat{f}_\lambda$ "best" predicts the original data points. If $\hat{f}_\lambda^{[k]}(t)$ denotes the smoothing spline estimate with the kth data point deleted, then $(\hat{f}_\lambda^{[k]}(t_k)-y_i)^2$ is a measure of the ability of $\hat{f}_\lambda^{[k]}(t_k)$ to predict y_k. Therefore, λ is chosen to minimize the cross-validation function

$$CV(\lambda)=\frac{1}{n}\sum_{k=1}^{n}(y_k-\hat{f}_\lambda^{[k]}(t_k))^2.$$

Craven and Wahba (1979) have shown that by weighting this sum correctly, one can define a ***generalized cross-validation function***, $GCV(\lambda)$, that has some nice properties. They have shown that if

$$GCV(\lambda)=\frac{1}{n}\sum_{k=1}^{n}(y_k-\hat{f}_\lambda^{[k]}(t_k))^2\, w_k(\lambda)$$

where $w_k(\lambda)=[(1-a_{kk}(\lambda))/(\frac{1}{n}tr(I-A(\lambda))]^2$ and $a_{kk}(\lambda)$ is the kth diagonal element of $A(\lambda)$, then under some general conditions

$$E(GCV(\lambda))\approx E(ASE(\lambda))+\sigma^2$$

for λ in the neighborhood of the minimizer of $E(ASE(\lambda))$. Therefore, a good estimate of $\hat{\lambda}_0$ which does not require knowledge of σ^2 is the minimizer of $GCV(\lambda)$, denoted $\hat{\lambda}_{GCV}$. Other desirable properties of $GCV(\lambda)$ and $\hat{\lambda}_{GCV}$ have been shown in Craven and Wahba (1979), Speckman (1983), Rice (1984) and Li (1986).

3.3 Estimating the derivative

Probably the most basic way to estimate $f'(t)=\frac{d}{dt}f(t)$ from the noisy data (t_i, y_i) where

$$y_i = f(t_i)+e_i \qquad i=1,\ldots,n$$

is to compute the derivative as $(y_i-y_{i-1})/(t_i-t_{i-1})$. However, there have been several other more sophisticated numerical methods proposed in the literature that seem to have much better properties. For example, Cullin (1971) suggests an approach that has been derived using the fact that a derivative can be written as the solution to a Fredholm integral equation of the first kind, and this equation is regularizable. Closely related to this approach is a time series approach taken by Anderssen and Bloomfield (1974a,b). More recently, estimates of derivatives have been derived using nonparametric regression methods. A kernel estimate approach is suggested by Schuster and Yakowitz (1979) and by Gasser and Müller (1984). A smoothing spline approach is suggested and discussed in Reinsch (1967), Wold (1974), Wahba (1975), Rice and Rosenblatt (1981) and Silverman (1985).

One possible estimate of f' using kernel methods is

$$\hat{f}_b{}'(x)=\frac{d}{dx}\hat{f}_b(x)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b^2}K'\left(\frac{x-t_i}{b}\right)y_i\,.$$

where $K'(b^{-1}(x-t_i))=\frac{d}{dx}K(b^{-1}(x-t_i))$. Gasser, Müller and Mammitzsch (1985) show that using K' to estimate f' is not necessarily optimal and they suggest an alternate kernel. However, we will use the definition above because it is analogous to the smoothing spline estimate of a derivative. Härdle, Marron and Wand (1988) suggest that there is not a great deal of practical difference between using K' and using the optimal kernel. They also address the problem of how to choose b when estimating a derivative.

A different but analogous estimate of f' using smoothing splines is $\hat{f}_\lambda{}'(x)=\frac{d}{dx}\hat{f}_\lambda(x)$

where $\hat{f}_\lambda(x)$ is an $(m+1)$th order smoothing spline estimate of f. We can justify this estimate by viewing smoothing splines in a larger context. Given data that follow the model

$$y_i=(Sf)(t_i)+\epsilon_i$$

where S is a linear operator, an estimate of f is the function $g\in W_2^m[0,1]$ that minimizes

$$\frac{1}{n}\sum_{i=1}^{n}(y_i-(Sg)(t_i))^2+\lambda\int_{[0,1]}(g^{(m)}(u))^2du$$

(Rice and Rosenblatt, 1983). In the case of estimating a derivative, S is defined to be the operator such that $(Sf)(t)=\int_{[0,t]} f(s)ds$. Using this definition of S, if we define functionals L_1 and L_2 where

$$L_1(g)=\frac{1}{n}\sum_{i=1}^{n}(y_i-(Sg)(t_i))^2+\lambda\int_{[0,1]}(g^{(m)}(u))^2du$$

and

$$L_2(f)=\frac{1}{n}\sum_{i=1}^{n}(y_i+f(0)-f(t_i))^2+\lambda\int_{[0,1]}(f^{(m+1)}(u))^2du$$

then I have been able to show that

$$\frac{d}{dt}\min_{f\in\mathcal{H}_{m+1}}L_2(f)=\min_{g\in\mathcal{H}_m}L_1(g)$$

where $\mathcal{H}_m=W_2^m[0,1]$ and $\mathcal{H}_{m+1}=W_2^{m+1}[0,1]$. Note that if $f\in W_2^{m+1}[0,1]$, then f is absolutely continuous so that we can write $f(t)-f(0)=\int_{[0,t]} g(s)ds$ where $g=f'$, and hence $g\in W_2^m[0,1]$. The problem of how to choose λ when minimizing L_1 is discussed in Rice (1986).

4. SMOOTHING SPLINE ESTIMATE OF THE RATE EQUATION Φ

In this section, we examine the statistical properties of an estimate of the function $\Phi(u)=f'(f^{-1}(u))$ where f is strictly decreasing. The proposed estimate takes the form $\hat{\Phi}_\lambda(u)=\hat{f}'_\lambda(\hat{f}_\lambda^{-1}(u))$ where $\hat{f}_\lambda(t)$ is a smoothing spline estimate, constrained to be strictly decreasing, λ is the smoothing parameter, $\hat{f}'_\lambda(t)=\frac{d}{dt}\hat{f}_\lambda(t)$ and $\hat{f}_\lambda^{-1} \circ \hat{f}_\lambda = I = \hat{f}_\lambda \circ \hat{f}_\lambda^{-1}$ where I is the identity mapping. Villalobos and Wahba (1987) and Kelly and Rice (1988) discuss constrained smoothing spline estimates.

Recall that in our application, the function Φ relates nitrogen concentration of a solution surrounding corn plant roots to uptake rate of nitrogen through corn plant roots. The function f relates time elapsed since corn plant roots are placed in solution to nitrogen concentration of the surrounding solution. Using the data from Figure 1a, I have displayed the (unconstrained) smoothing spline estimate of f in Figure 2a and the corresponding estimate of the rate curve Φ in Figure 2b. In Figures 3a and 3b, I have diplayed just the smoothing spline rate curve estimates from eight different, but related data sets.

As mentioned earlier, studying the theoretical properties of kernel estimators is more tractable than studying the theoretical properties of smoothing splines. Therefore, in the remainder of this section I examine some properties of the estimator $\hat{\Phi}_b(u)=\hat{f}'_b(\hat{f}_b^{-1}(u))$ where $\hat{f}_b(t)$ is a kernel estimate, constrained to be strictly decreasing, b is the smoothing parameter, $\hat{f}'_b(t)=\frac{d}{dt}\hat{f}_b(t)$ and $\hat{f}_b^{-1} \circ \hat{f}_b = I = \hat{f}_b \circ \hat{f}_b^{-1}$ where I is the identity mapping. Using the results of Nychka (1989), these results should extend to smoothing splines.

4.1 Consistency of an estimate of Φ

In this section I will show that $\hat{\Phi}_b$ is a consistent estimate of Φ. To do so, it is helpful to introduce some notation. Denote the domain of f as $\mathcal{T}$ and the range of f as $\mathcal{C}$ so that $f:\mathcal{T}\rightarrow\mathcal{C}$ and $f^{-1}:\mathcal{C}\rightarrow\mathcal{T}$ are bijective. In addition, define the following norm

$$\|f\|_{\mathcal{A}}=\sup_{x\in\mathcal{A}}|f(x)|$$

which has the property that

$$\|f\cdot g\|_{\mathcal{A}} \leq \|f\|_{\mathcal{A}} \|g\|_{\mathcal{A}}. \tag{*}$$

Note that since f and f^{-1} are bijective, we have the relation

$$\|f^{-1}\|_{\mathcal{C}} = \|f^{-1} \circ f\|_{\mathcal{T}} \tag{**}$$

since, if c=f(t) then

$$\sup_{c\in\mathcal{C}} |f^{-1}(c)| = \sup_{t\in\mathcal{T}} |f^{-1}(f(t))|.$$

Assumption A Let $\hat{f}_b(t)$ be a kernel estimate, constrained to be strictly decreasing and let $\hat{f}_b'(t) = \frac{d}{dt}\hat{f}_b(t)$. Then as $n\to\infty$, we have $b=b(n)\to 0$, $nb\to\infty$, $\|\hat{f}_b - f\|_{\mathcal{T}} = o_P(1)$ and $\|\hat{f}_b' - f'\|_{\mathcal{T}} = o_P(1)$.

Assumption A is not unreasonable for nonparametric regression estimates. In fact, Cox (1988) has verified for smoothing spline estimates that under some reasonable conditions, if $b=b(n)\to 0$ and $nb\to\infty$ as $n\to\infty$, then $\|\hat{f}_\lambda - f\|_{\mathcal{T}} = o_P(1)$ as $n\to\infty$. The reader is referred to this same paper for a general discussion of function spaces and norms natural to the smoothing spline problem. Next, a consistency result for $\hat{\Phi}_b$ is stated.

Theorem *Under Assumption A, and assuming that* $M_{n,b} = O_P(1)$ *where* $M_{n,b} = \|\hat{f}_b''\|_{\mathcal{T}} \|1/\hat{f}_b'\|_{\mathcal{T}}$, *we have*

$$\|\hat{\Phi}_b - \Phi\|_{\mathcal{C}} \to 0 \text{ \textit{in probability}}$$

as $n\to\infty$.

Proof For clarity, we define $\hat{g}_b = \hat{f}_b^{-1}$. We proceed by using the triangle inequality and then by applying (**).

$$\|\hat{\Phi}_b - \Phi\|_{\mathcal{C}} = \|\hat{f}_b' \circ \hat{g}_b - f' \circ f^{-1}\|_{\mathcal{C}}$$

$$\leq \|\hat{f}_b' \circ \hat{g}_b - \hat{f}_b' \circ f^{-1}\|_{\mathcal{C}} + \|\hat{f}_b' \circ f^{-1} - f' \circ f^{-1}\|_{\mathcal{C}}$$

$$= \|\hat{f}_b' \circ \hat{g}_b \circ f - \hat{f}_b' \circ f^{-1} \circ f\|_{\mathcal{T}} + \|\hat{f}_b' - f'\|_{\mathcal{T}}$$

Consider the first term on the right hand side. By applying the Mean Value Theorem to $\hat{f}_b'$ and then applying (*), we have

$$\|\hat{f}_b' \circ \hat{g}_b \circ f - \hat{f}_b' \circ f^{-1} \circ f\|_{\mathcal{T}} \leq \|\hat{f}_b''\|_{\mathcal{T}} \, \|\hat{g}_b \circ f - I\|_{\mathcal{T}}$$

If we rewrite I, the identity operator, as $\hat{g}_b \circ \hat{f}_b$, (recall that $\hat{g}_b = \hat{f}_b^{-1}$), apply the Mean Value Theorem to $\hat{g}_b$ and again apply (*), this last expression becomes

$$\|\hat{f}_b''\|_{\mathcal{T}} \, \|\hat{g}_b \circ f - \hat{g}_b \circ \hat{f}_b\|_{\mathcal{T}} \leq \|\hat{f}_b''\|_{\mathcal{T}} \, \|\hat{g}_b'\|_{\mathcal{T}} \, \|f - \hat{f}_b\|_{\mathcal{T}}$$

$$= \|\hat{f}_b''\|_{\mathcal{T}} \, \|1/\hat{f}_b'\|_{\mathcal{T}} \, \|f - \hat{f}_b\|_{\mathcal{T}}$$

noting that $\hat{g}_b' = \frac{d}{dc} \hat{f}_b^{-1}(c) = \frac{1}{\hat{f}_b'(t)}$ where $c = f(t)$. Putting everything together, we have

$$\|\hat{\Phi}_b - \Phi\|_{\mathcal{C}} \leq \|\hat{f}_b''\|_{\mathcal{T}} \, \|1/\hat{f}_b'\|_{\mathcal{T}} \, \|f - \hat{f}_b\|_{\mathcal{T}} + \|\hat{f}_b' - f'\|_{\mathcal{T}}$$

$$= M_{n,b} \, \|f - \hat{f}_b\|_{\mathcal{T}} + \|\hat{f}_b' - f'\|_{\mathcal{T}}$$

$$\rightarrow 0 \text{ in pr}$$

where $M_{n,b}=\|\hat{f}_b''\|_{\mathcal{T}}\|1/\hat{f}_b'\|_{\mathcal{T}}$. □

Note that this rate of convergence is dependent on the size of the second derivative and on one over the first derivative. Therefore, we do not want f to be very wiggly. Also if a portion of f is nearly flat then f' will be very small and the rate of convergence will be slow. This may be the case with the right boundary of a depletion curve. Therefore, one might want to use exponential smoothing near the boundaries of f. See Rice and Rosenblatt (1981) and Eubank (1988) for discussion on boundary effects of nonparametric regression methods.

4.2 Estimating Φ when a parametric form, Φ_θ is specified

The conventional approach to estimating Φ first involves specifying some parametric form for Φ, say $\Phi_{\underset{\sim}{\theta}}$. If the first order differential equation $\Phi_{\underset{\sim}{\theta}}(C)=-C'$ where $C(0)=c_0$ has an explicit closed form solution, then we denote that solution as $C_{\underset{\sim}{\theta}}(t)$. Then, using $C_{\underset{\sim}{\theta}}(t)$ as the model for the depletion curve C(t), $\underset{\sim}{\theta}$ can be estimated directly from the data (t_i, y_i) using standard linear or nonlinear regression techniques. However, when there is no explicit closed form solution, the time derivative of C must be estimated from the data, and then Φ must be estimated using these estimates.

Historically, it has been assumed that the uptake curve Φ follows the Michealis-Menten model

$$-\Phi_{\underset{\sim}{\theta}}(u) = \frac{V_{max}\cdot u}{K_m+u}$$

where $\underset{\sim}{\theta}^T=[V_{max}, K_m]$. Note that there is no closed form solution to the system $-C'=a\cdot C/(b+C)$ where $C(0)=c_0$. The parameters in the Michealis-Menten model have a nice biological interpretation: V_{max} is the maximum uptake rate that is approached as the concentration increases; K_m is the concentration that yields an uptake rate equal to $V_{max}/2$. In the past, these two parameters have been used to completely describe an uptake curve and to compare two different uptake curves. However, work by Jackson and others give evidence indicating that the Michealis-Menten model does not adequately describe an uptake curve. This might also be suggested in Figures 3a and 3b. (For those applications where the Michealis-Menten model is adequate and its variables can be measured directly, Ruppert, Cressie and Carroll (1988) discuss how to estimate its parameters).

In the remainder of this section, we assume that the parametric model $\Phi_{\underset{\sim}{\theta}}$ adequately describes the uptake curve and that there is no closed form solution to the system $\Phi_{\underset{\sim}{\theta}}(C) = -C'$, $C(0) = c_0$. Conventionally, $\underset{\sim}{\theta}$ is estimated by $\hat{\underset{\sim}{\theta}}_D$ defined as

$$\hat{\underset{\sim}{\theta}}_D = \min_{\underset{\sim}{\theta}} \frac{1}{n} \sum_{i=1}^{n} [r_i - \Phi(y_i; \underset{\sim}{\theta})]^2 \, w_i$$

where $r_i = (y_i - y_{i-1})/(t_i - t_{i-1})$ and $w_i \propto \mathrm{var}(y_i)$. I use the subscript D to indicate that this estimate uses a derivative estimate based on *differenced* data. Note that this is just ordinary weighted least squares.

My proposal is to estimate $\underset{\sim}{\theta}$ by $\hat{\underset{\sim}{\theta}}_S$ defined as

$$\hat{\underset{\sim}{\theta}}_S = \min_{\underset{\sim}{\theta}} \int_{[0,1]} [\hat{C}_b'(t) - \Phi(\hat{C}_b(t); \underset{\sim}{\theta})]^2 \, w(\hat{C}_b(t)) \, \hat{C}_b'(t) \, dt$$

where $\hat{C}_b(t)$ is a constrained kernel estimate of the depletion curve C, $\hat{C}_b'(t) = \frac{d}{dt}\hat{C}_b(t)$ and $w(\cdot)$ is a smooth weight function. I use the subscript S to indicate that this estimate uses a derivative estimate based on *smoothed* data. I arrive at $\hat{\underset{\sim}{\theta}}_S$ by minimizing $\|\hat{\Phi} - \Phi\|^2$ where the norm is a weighted L_2 norm and then by doing a change of variable on concentration $u = \hat{C}_b(t)$. Under some reasonable boundedness assumptions, I have been able to show that

$|\hat{\theta}_S - \theta|/b_n \xrightarrow{\mathfrak{D}} N(0,1)$ as $n \to \infty$ where $b_n \to 0$. This implies that $\hat{\theta}_S$ is a consistent estimate of θ (Serfling, 1980). My conjecture is that this result can be extended to an arbitrary nonlinear model for $\Phi_{\underset{\sim}{\theta}}$. In addition, I conjecture that $\hat{\theta}_D$ is not consistent.

5. SUGGESTED FUTURE WORK

The majority of this paper gives background information necessary for defining the problem of interest, and that is, to obtain a good estimate of the rate equation Φ. I have shown just a few preliminay results in section 4. The first thing I need to work on is extending these results from the kernel estimate case to the smoothing spline estimate case. From there, I propose to investigate several of the follow areas:

1. Extend the results in section 4.2 to the case where $\Phi(u;\, \underset{\sim}{\theta}) = \frac{u \cdot \theta_1}{\theta_2 + u}$ and then to the case where $\Phi(u;\, \underset{\sim}{\theta})$ an "arbitrary" nonlinear model.

2. Determine a how to choose the optimal smoothing parameter for $\hat{\Phi}_\lambda$.

3. Determine how to make boundary adjustments to $\hat{\Phi}_\lambda$.

4. Derive estimates of certain features of Φ and give confidence bounds for these estimates.

5. Construct a goodness-of-fit type test to determine whether a parametric form $\Phi(u;\, \theta)$ is reasonable.

6. Determine how to optimally choose the n time points in a solution depletion experiment. Two references are Micchhelli and Wahba (1981) and Wahba (1971).

6. REFERENCES

Anderssen, R. S. and Bloomfield, P. (1974a). A time series approach to numerical differentiation. *Technometrics* 16, 69-75.

Anderssen, R. S. and Bloomfield, P. (1974b). Numerical differentiation procedures for non-exact data. *Numerische Mathematik* 22, 157-182.

Cox, D. D. (1984). Multivariate smoothing spline functions. *SIAM Journal of Numerical Analysis* 21, 789-813.

Cox, D. D. (1988). Approximation of the method of regularization estimates. *Annals of Statistics* 16, 694-712.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-vaslidation. *Numerische Mathematik* 31, 377-403.

Cullum, J. (1971). Numerical differentiation and regularization. *SIAM Journal of Numerical Analysis* 8, 254-265.

de Boor, G. (1972). On calculating with B-splines. *Journal of Approximation Theory* 6, 50-62.

Eubank, R. L. (1988). *Spline smoothing and nonparametric regression.* New York: Marcel Dekker, Inc.

Gasser, T. and Müller, H. G. (1979). Nonparametric estimation of regression functions and their derivatives. Preprint No. 38, Sonderforschungsbereich 123, Universität Heidelberg.

Gasser, T. and Müller, H. G. (1984). Estimating regression functions and their derivatives by the kernel method. *Scandanavian Journal of Statistics* 11, 197-211.

Gasser, T., Müller, H. G. and Mammitzsch, V. (1985). Kernels for nonparametric curve estimation. *Journal of the Royal Statistics Society*, Series B 47, 238-252.

Härdle, W., Hall, P. and Marron, J. S. (1988). How far are automatically chosen regression smoothing parameters from their optimum? *Journal of the American Statistical Association* 83, 86-95.

Härdle, W., Marron, J. S. and Wand M. P. (1988). Bandwidth choice for density derivatives. *Journal of the Royal Statistics Society*, Series B. To appear.

Jackson, W. (1989). Personal communication.

Keeton, W. T. (1980). *Biological sciences*, Third edition. New York: W. W. Norton and Company.

Kelly, C. and Rice, J. (1988). Monotone smoothing with application to dose response curves and the assessment of synergism. Unpublished manuscript.

Li, K. C. (1986). Asymptotic optimality of C_L and generalized cross-validation in ridge regression with application to spline smoothing. *The Annals of Statistics* 14, 1101-1112.

Mallows, C. L. (1973). Some comments on C_p. *Technometrics* 15, 661-675.

Micchhelli, C. A. and Wahba, G. (1981). Design problems for optimal surface interpolation. *Approximation Theory and Applications*, 329-347.

Müller, H. G. (1988). *Nonparametric regression analysis of longitudinal data* (*Lecture Notes in Statistics*, 46). Heidelberg: Springer-Verlag.

Nadaryra, E. A. (1964). On estimating regression. *Theory of Probability and its Applications* 9, 141-142.

Nychka, D. (1988). Bayesian confidence intervals for smoothing splines. *Journal of the American Statistical Association* 83, 1134-1143.

Nychka, D. (1989). Smoothing splines as locally weighted averages. Mimeo. Series No. 1960, Institute of Statistics, North Carolina State University.

O'Sullivan, F. (1986). A statistical perspective on ill-posed inverse problems. *Statistical* splines. Unpublished manuscript.

Villalobos, M. and Wahba, G. (1987). Inequality-constrained multivariate smoothing splines with application to the estimation of posterior probabilities. *Journal of the American Statistical Society* 82, 239-248.

Wahba, G. (1971). On the regression design problem of Sacks and Ylvisaker. *The Annals of Mathematical Statistics* 42, 1035-1053.

Wahba, G. (1975). Smoothing noisy data with spline functions. *Numerische Mathematik* 24, 383-393.

Wahba, G. (1983). Bayesian "confidence intervals" for the cross-validated smoothing spline. *Journal of the Royal Statistical Society*, Series B 45, 133-150.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā A* 26, 359-372.

Whittaker, E. T. (1923). On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society* 41, 63-75.

Science 1, 502-527.

Priestly, M. B. and Chao, M. T. (1972). Non-parametric function fitting. *Journal of the Royal Statistical Society*, Series B 34, 385-392.

Rice, J. (1984). Bandwidth choice for nonparametric regression. *The Annals of Statistics* 12, 1215-1230.

Rice, J. (1986). Bandwidth choice for differentiation. *Journal of Multivariate Analysis* 19, 251-264.

Rice, J. and Rosenblatt, M. (1981). Integrated mean squared error of a smoothing spline. *Journal of Approximation Theory* 33, 353-369.

Rice, J. and Rosenblatt, M. (1983). Smoothing splines: Regression, derivatives and deconvolution. *the Annals of Statistics* 11, 141-156.

Ruppert, D., Cressie, N. and Carroll, R. (1988). A transformation/weighting model for estimating Michaelis-Menten parameters. Technical Report No. 796, School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca.

Schoenberg, I. J. (1964). Spline functions and the problem of graduation. *Proceeds of the National Academy of Science USA* 52, 947-950.

Schuster, E. and Yakowitz, S. (1979). Contribution to the theory of nonparametric regression, with application to system identification. *The Annals of Statistics* 7, 139-149.

Serfling, S. R. (1980). *Approximation Theorems of Mathematical Statistics.* New York: John Wiley.

Silverman, B. W. (1985). Some aspects of the spline smoothing approach to non-parametric regression curve fitting (with discussion). *Journal of the Royal Statistical Society*, Series B 47, 1-52.

Speckman, P. (1982). Efficient nonparametric regression with cross-validated smoothing

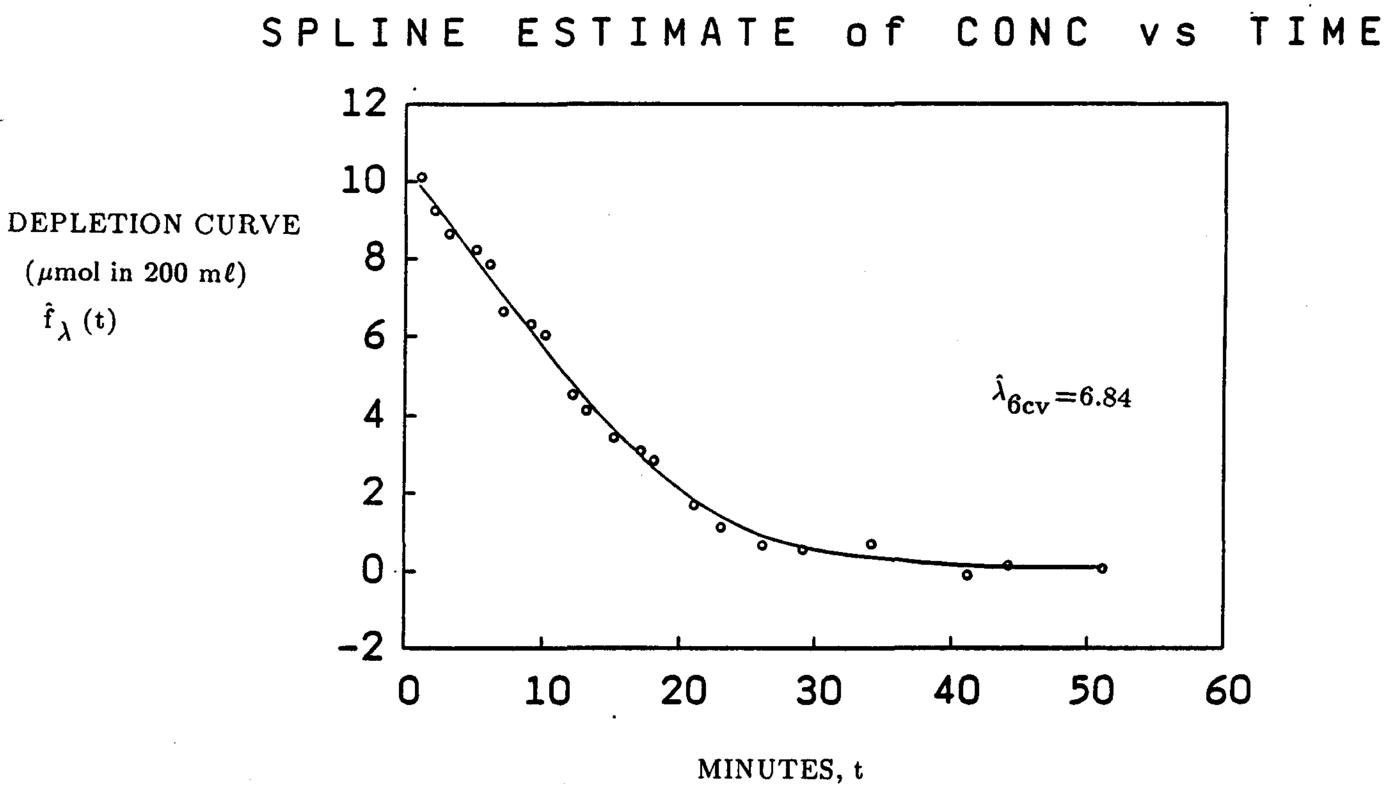
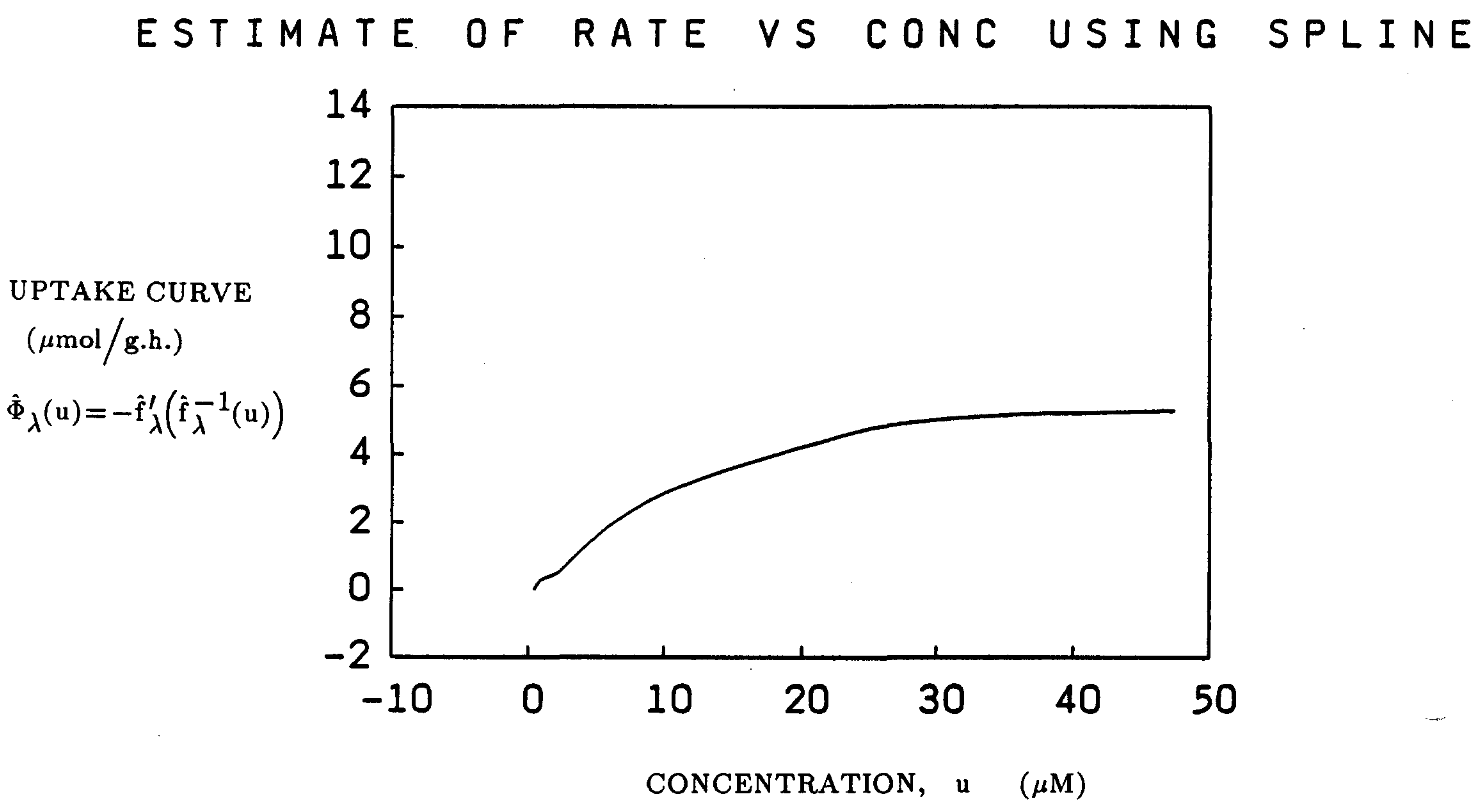


FIGURE 2a (above): Unconstrained smoothing spline estimate of the depletion curve f.

FIGURE 2b (below): Unconstrained smoothing spline estimate of the uptake curve Φ.



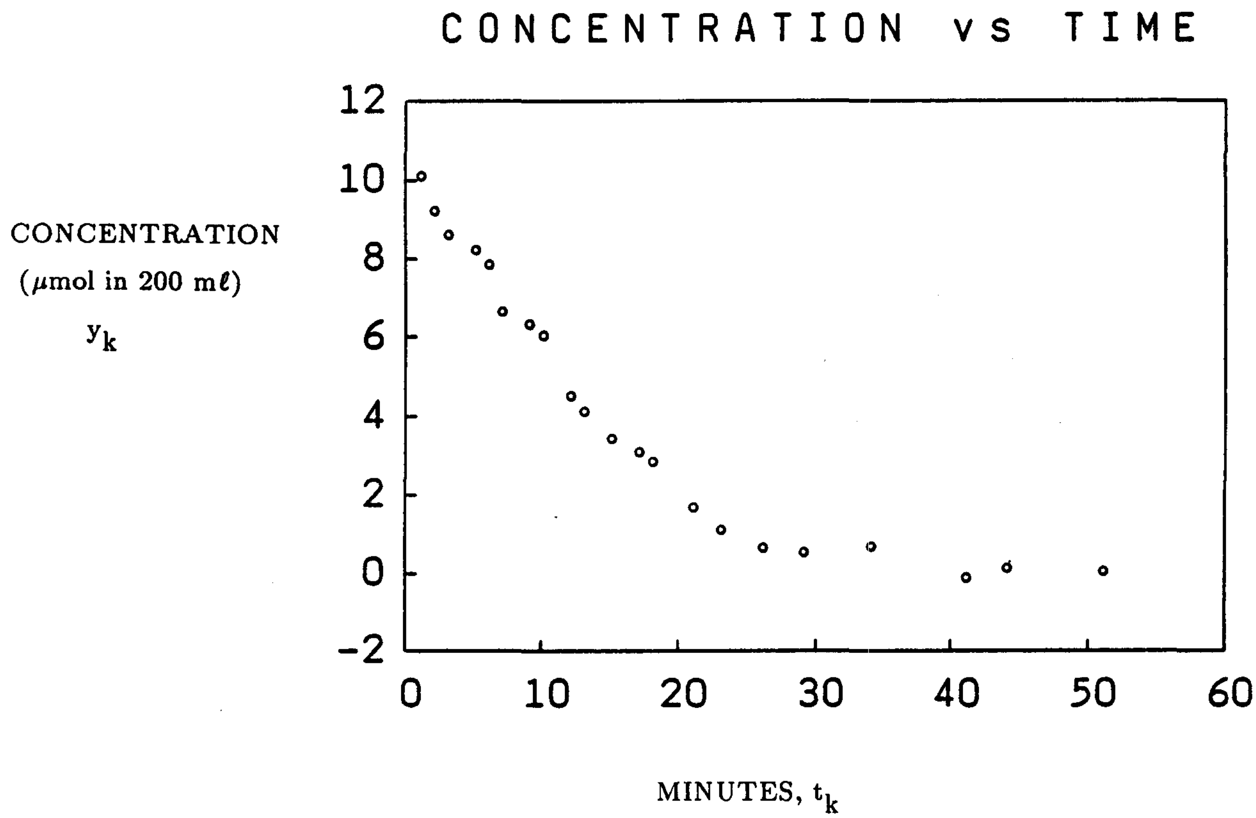
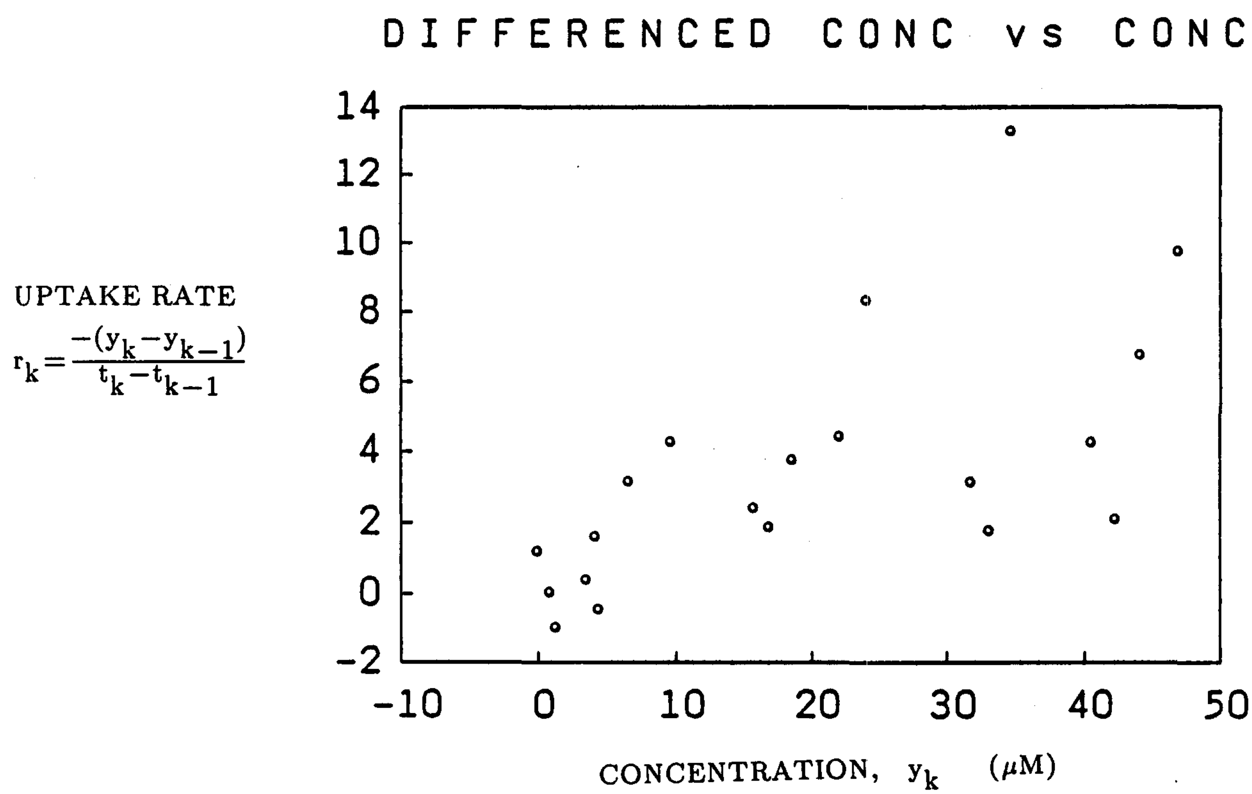
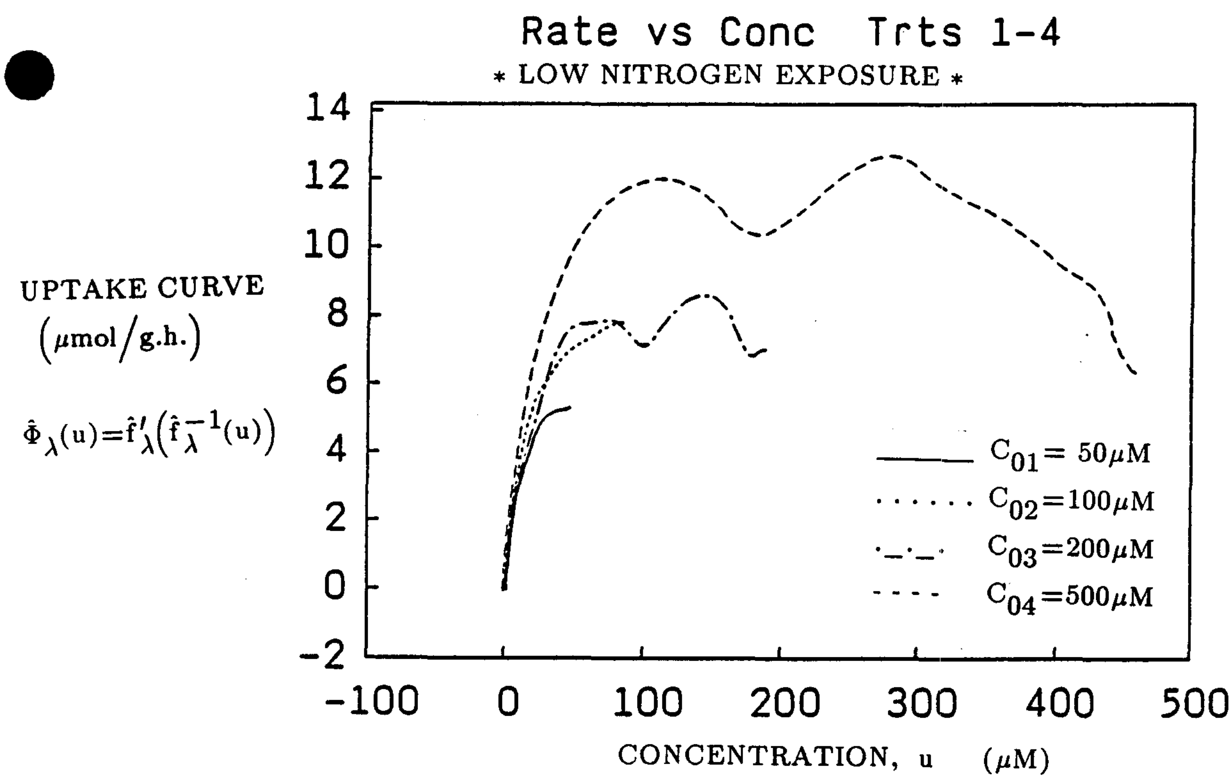


FIGURE 1a (above): Data from a solution depletion experiment with $c_o = 10\mu\text{mol}/200\text{m}\ell = 50\mu\text{M}$

FIGURE 1b (below): Pointwise derivative estimates using first differences of data in Figure 1a.





FIGURES 3a (above) and 3b (below): Unconstrained smoothing spline estimates of rate curves from eight different solution depletion experiments. The eight comes from four different initial concentrations: C_{01}, C_{02}, C_{03}, C_{04}; and from two different levels of previous nitrogen exposure.