NONPARAMETRIC ESTIMATION OF RATE EQUATIONS FOR NUTRIENT UPTAKE

Kristen Meier and Douglas Nychka

Institute of Statistics Mimeograph Series No. 1995

July, 1991

NORTH CAROLINA STATE UNIVERSITY
Raleigh, North Carolina

NONPARAMETRIC ESTIMATION OF RATE EQUATIONS FOR NUTRIENT UPTAKE

by

Kristen Meier[1,2] and Douglas Nychka[2]

[1]Statistics and Biomathematics Branch
National Institute of Environmental Health Sciences, P.O. Box 12233
Research Triangle Park, North Carolina 27709

[2]Department of Statistics
North Carolina State University
Raleigh, North Carolina 27695-8203

July 10, 1991

ABSTRACT

The rate of a biological process is important for characterizing the system and is necessary for gaining a deeper understanding of the process. Consider measurements, Y, that are made over time on a system following the model $Y = f(t) + e$, where f is a smooth, unknown function and e is measurement error. While most statistical methodology has focused on estimating $f(t)$ or $f'(t)$, in some applications what is of real biological interest is the relationship between f and f'. One example is the study of nitrogen absorption by plant roots through a solution depletion experiment. In this case $f(t)$ is the nitrate concentration of the solution surrounding the roots at time t and $-f'(t)$ is the absorption rate of nitrate by plant roots at time t. One is interested in the rate of nitrate absortion as a function of concentration, that is, one is interested in Φ, where $\Phi(f) = -f'$. Knowledge of Φ is important in quantifying the ability of a particular plant species to absorb nitrogen and in comparing the absorption ability of different crop varieties. A parametric model for Φ is not usually available and thus a nonparametric estimate of Φ is particularly appropriate. This work proposes using spline based curve estimates with the smoothing parameter chosen by cross-validation. These methods are used to analyze a series of solution depletion experiments and are also examined by a simulation study designed to mimic the main features of such data. Although the true f is a monotonic function, simulation results indicate that constraining the estimate of f to be monotonic does not reduce the average squared error of the rate curve estimate, Φ. While using a cross-validated estimate of the smoothing parameter tends to inflate the average squared error of the rate estimate, an analysis of a set of solution depletion experiments is still possible. Using the proposed methods, we are able to detect a difference in rate curves obtained under different experimental conditions. This is established by an ANOVA-like test applied to the estimated rate curves where the critical value is determined by a parametric version of the bootstrap. This finding is important because it suggests that the shape of Φ may not be constant under the experimental conditions examined.

Key Words and Phrases Solution depletion experiments, Michaelis-Menten rate equation, smoothing splines, monotonicity constraints, cross-validation, bootstrap

Acknowledgments This application in soil science would not have been possible without the enthusiastic support of Dr. William Jackson, North Carolina State University. We would also like to thank Peggy Longmeier for her work on the experimental design and doing much of the laboratory work. This research was supported by the National Sciences Foundation Grant DMS-8715756 and the National Institute of Environmental Health Sciences.

1. INTRODUCTION

Nitrogen is an essential mineral for plant growth. Plants absorb nitrogen from the soil mainly in the form of nitrate or ammonium ions. However, at any given time, most soils contain only small amounts of these ions relative to the amount continuously required by plants (Epstein 1972). Nitrogen is added to the soil through the atmosphere (brought down by precipitation), through plant residues, through animal wastes and of increasing importance, through the application of nitrogenous fertilizers. However, only a small portion of this nitrogen is in the form of nitrate or ammonium. In addition, ammonium is easily converted to nitrate and nitrate is readily leached from the soil by rain or irrigation water, possibly becoming a water pollutant (Brady 1984). Therefore, knowledge of the nitrogen *uptake* (absorption) process by plant roots is necessary to avoid both nitrogen deficiency and leaching.

The rate that roots absorb nitrate ions is a function of the nitrate concentration of the medium surrounding the roots (Epstein 1972). The dependence of the uptake rate, r, on the ambient concentration, c, is assumed to follow a smooth function $\Phi(c) = r$. Knowledge of the functional form of Φ is important in quantifying the ability of a particular type of plant to absorb a specific ion. The mechanisms in roots cells that control the absorption of nitrate are not completely understood; identifying Φ is a first step in modeling the cellular processes that mediate nitrate transport.

It is difficult to measure uptake rates directly because the absorption process by the plant roots must invariably change the ambient concentration. One direct measure is to use a radioactively labeled ion and then assay the amount of labeled nutrient that is absorbed by the roots in a short period of time. Unfortunately, measuring nitrate absorption using this technique is difficult because the radioactive isotopes of nitrogen have short decay times. An alternative method is to determine Φ indirectly by a solution depletion experiment (Claassen and Barber 1974). A single plant or set of plants is placed in a beaker containing a specific volume and prespecified initial concentration, c_0, of an aerated nutrient solution. Then, at each of n predetermined time points, $t_1,t_2,...t_n$, the solution in the beaker is sampled and the concentrations, $y_1,y_2,...y_n$ are determined. Figure 1 illustrates typical results for solution depletion experiments that we have encountered for the uptake of NO_3^- (nitrate) by maize roots. The ability of the root material to deplete the solution of nitrate ions over time is readily apparent. However, the shape of the uptake rate curve, Φ, is not clear from these time plots. A statistical problem then is to recover reliable estimates of Φ from such depletion experiments.

It will be assumed that the depletion data follow the additive model

$$y_k = f(t_k) + e_k, \qquad 1 \leq k \leq n \quad \text{and} \quad 0 = t_1 < t_2 < \ldots < t_n = T \tag{1.1}$$

where the depletion curve, f, is a smooth function of time and $\{e_k\}$ are random, independent errors with constant variance σ^2. Typically, statistical methodology has focused on estimating f or its derivative from $\{(t_k, y_k)\}$. However, in this application and in many other biological applications, what is of real interest is not f(t) or $f'(t)$ directly, but the *relationship* between f and f'. That is, one is interested in estimating the function Φ, where $\Phi(f) = -f'$.

Because the biochemical basis of nitrate absorption has not been completely identified, it is difficult to specify parametric models for Φ. Also, because depletion experiments do not measure the uptake rate directly, it is not possible to readily identify empirical models for Φ based on a graphical examination of the data. For these reasons, it is appropriate to consider nonparametric regression estimates of Φ. While this nonparametric approach may not give direct insight into the internal mechanisms of ion absorption, it does enable us to quantify how rates of absorption are affected by changes in various external factors or internal conditions. In this article we apply smoothing spline methods to estimate Φ and suggest a simulation-based procedure to test for differences among rate curves for different treatment groups.

Although much of this article focuses on a specific application of nonparametric regression in soil science, there are two issues of more general interest. These are (1) the advantages of monotonicity constraints on the estimated depletion curve and (2) the data-based selection of the smoothing parameter (bandwidth). Basic scientific considerations imply that f should be a monotonically decreasing function in time. However, with most nonparametric estimators it is difficult to enforce monotonicity. We are interested in quantifying (in the context of our application) the benefit of using monotonic estimators when the true function is indeed monotonic. The accuracy of the rate curve estimate depends primarily on the quality of the estimate of f'. Most data-based methods of smoothing parameter selection attempt to minimize the expected average squared error with respect to f. One important question is whether smoothing parameter estimates based on cross-validation for the observed data are also appropriate for estimates of Φ.

The next section reviews some existing approaches to estimate rate curves and introduces an estimate of Φ based on smoothing splines. Section 3 gives the details of a particular depletion experiment designed to examine whether the uptake mechanism of nitrate ions in maize have "memory." That is, does the form of Φ depend on the initial concentration of nitrate in the depletion experiment? The distinction between constrained and unconstrained depletion curve estimates were investigated by a simulation study and these results are reported in Section 4. The last section draws some conclusions from the analysis of the depletion experiment and considers more general statistical issues associated with the use of spline estimators.

2. RATE CURVE ESTIMATORS

2.1 Parametric Models

Epstein and Hagen (1952) first proposed that the Michaelis-Menten model,

$$\Phi(c;\underset{\sim}{\theta}) = \frac{\theta_1 \cdot c}{\theta_2 + c},$$

be used to describe ion uptake as a function of concentration. Michaelis and Menten (1913) originally derived this model from enzyme reaction equations; Epstein and Hagen (1952) noted that the mechanism of ion absorption and enzyme activity are "quite analogous" and therefore the Michaelis-Menten model was appropriate. Later, extensions of the Michaelis-Menten model and combinations of several Michaelis-Menten models were proposed (see Epstein 1972). At present, it has been suggested that the uptake of ions, in particular nitrate, by plant roots is much more complicated than what is suggested by these earlier models (Clarkson 1986, Jackson, Volk, Morgan, Pan and Teyker 1986, and Jackson, Pan, Moll and Kamprath 1986). We attempted to model these suggested, more complicated mechanisms of ion absorption using a compartmental model (Jacquez 1985). However, at this point in time, obtaining data for such a model is not yet feasible (Meier 1990 and Jackson, personal communication). Therefore, this modeling approach was abandoned in favor of the totally nonparametric approach proposed in this paper.

There are computational difficulties associated with estimating parameters for the rate function, even if an adequate parametric model is available. In order to relate Φ to the observational data from a depletion experiment, it is necessary to solve the first order differential equation: $\Phi(f; \underset{\sim}{\theta}) = -f'$ for f, and then to compare the solution, f, to the observed data. In most cases, such as Michaelis-Menten kinetics, this equation cannot be solved explicitly in f and must be solved numerically. Claassen and Barber (1974) first proposed fitting the Michaelis-Menten model to solution depletion data using this approach. They also suggest a nonparametric approach that is described in the next section.

2.2 Nonparametric Curve Estimates

If the depletion curve were known exactly, then the rate curve could be recovered by the relationship: $\Phi(c) = f' \circ g(c)$ where $g = f^{-1}$. We propose using a cubic smoothing spline to estimate f and then differentiate this estimate for the derivative estimate. The pairs of points $(\hat{f}(t), \hat{f}'(t))$ can be interpreted as estimates of $(c, \Phi(c))$ with $c = f(t)$. If $\hat{f}$ is monotonic then $\hat{\Phi}$ is actually a well defined function with respect to c and can be written explicitly as :

$$\hat{\Phi}(c) = \hat{f}'(\hat{f}^{-1}(c)). \tag{2.1}$$

Otherwise it is not possible to express this curve as a function across the entire range of observed concentrations. Several different methods have been proposed for estimating derivatives from noisy data (Reinsch 1967, Cullum 1971, Anderssen and Bloomfield 1974a,b, Wahba 1975, Rice and Rosenblatt 1983, Gasser and Müller 1984, Silverman 1985, Rice 1986). However, we are not aware of any published work that estimates a derivative of a function as it relates to the function itself (versus time) using the techniques proposed in this paper. Claassen and Barber (1974) suggest obtaining point estimates of the rate curve at the observed concentrations by fitting a parabolic spline function (DuChateau, Nofziger, Ahuja, and Swartzendruber 1972) to solution depletion data or by fitting a cubic spline to the data with variable knot locations. They differentiate the splines to obtain point estimates of the rate at each observed concentration. These researchers do not address the problems of selecting the number of knots in the spline estimate nor how one might draw statistical inferences from the estimated rates.

While our approach is similar to the nonparametric strategy suggested by Claassen and Barber, we have added several improvements to the nonparametric estimator. Most notably we consider monotonic estimates of the depletion curve and determine the amount of smoothing objectively by cross-validation. The details of our methods are given in the next subsection.

2.3 Smoothing Splines and Monotonicity Constraints

Although splines are popularly associated with piecewise polynomial functions, a more fundamental definition of a spline is as a solution to a variational problem. Let $\mathcal{H}=W_2^2[0,1]=\{h\colon h, h'$ absolutely continuous, $h'' \in L^2[0,1]\}$. For $\lambda > 0$, a cubic smoothing spline estimate (with natural boundary conditions) of the regression function f under the model (1.1), is the $h \in \mathcal{H}$ that minimizes

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

The minimizer of $\mathcal{L}(h)$ will exist and is unique provided that $n > 3$ and will be denoted as $\hat{f}_\lambda$ to emphasize its dependence on the smoothing parameter, λ. It is well known that $\hat{f}_\lambda$ can be written as a piecewise cubic polynomial with join points (knots) at each data point (Eubank 1988, Wahba 1990). For a fixed λ, $\hat{f}_\lambda$ is a linear function of the data, $\underset{\sim}{y}$, and one can define an $n \times n$ "hat" matrix $A(\lambda)$ such that $\hat{f}_\lambda(\underset{\sim}{t}) = A(\lambda)\underset{\sim}{y}$ where $\hat{f}_\lambda(\underset{\sim}{t}) = [\hat{f}_\lambda(t_1), \hat{f}_\lambda(t_2), \ldots, \hat{f}_\lambda(t_n)]^T$. Smoothing splines are one of several types of nonparametric regression estimates that are linear functions of $\underset{\sim}{y}$ and are essentially locally weighted averages (Silverman 1984, Nychka 1989). Here the smoothing parameter plays a similar role to the bandwidth for a kernel regression estimate, and an approximate expression for the average bandwidth of a cubic smoothing spline is $\lambda^{1/4}$.

The spline estimate described above may not be a monotonically decreasing function. This constraint can be imposed formally by simply restricting the minimization of $\mathcal{L}$ to $h \in \mathcal{H}$ such that $h' \leq 0$. Such estimators have been studied by Utreras (1985) and Villalobos and Wahba (1987) but are not readily computable. As an alternative, we choose to implement this constraint by exploiting the representation of the cubic smoothing spline as a linear combination of B-spline basis functions (Kelly and Rice 1990 , Ramsay 1988). From the characterization of a cubic smoothing spline estimate as a piecewise polynomial it is possible to represent the solution as a linear combination of cubic B-spline basis functions, $\{B_k\}$ (see de Boor (1978) for their construction). Moreover, it is easy to construct monotonic functions with respect to this B-spline basis. Suppose that

$$h(t) = \sum_{k=1}^{n} c_k B_k(t) \quad \text{and} \quad c_1 \geq c_2 \geq \ldots \geq c_n, \tag{2.2}$$

then h will be monotonically decreasing. This property suggests a monotonic spline, $\hat{f}_{M\lambda}$, that is the minimizer of $\mathcal{L}(h)$ for all h satisfying the conditions at (2.2). The main advantage of this representation is that the inequality constraints are on the coefficients rather than the function itself. This minimization problem is readily solved as it is a quadratic programing problem with inequality constraints (see Kelly and Rice 1990). Our work uses FORTRAN code kindly provided by John Rice and is based on de Boor's algorithm for evaluating B-splines (de Boor, 1972) and on Lawson and Hanson's (1974, Chapter 23) algorithm for computing constrained least squares estimates.

The spline estimate is sensitive to the choice of the smoothing parameter. Although λ is often chosen subjectively in the process of data analysis, we also recommend considering a data-based estimate of the smoothing parameter based on cross-validation. Let $\hat{f}_\lambda^{[k]}$ be the spline estimate of f having omitted the k^{th} data point and define the adjusted (for endpoints) cross-validation function as

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

$\mathrm{AdCV}(\lambda)$ is an estimate of the expected average prediction error of the spline for a particular choice of λ. Thus an estimate of λ is obtained by minimizing AdCV. For unconstrained splines, $y_k - \hat{f}_\lambda^{[k]}(t_k)$ can be shown be to equal to $(y_k - \hat{f}_\lambda(t_k))/(1 - A_{kk}(\lambda))$; therefore, if $\hat{f}_\lambda$ and the diagonal elements of $A(\lambda)$ are known, evaluating $\mathrm{AdCV}(\lambda)$ is only an order n calculation. This estimate for the optimal (with respect to expected average squared error) value of λ is also applicable to the monotonic smoothing spline described above; unfortunately the computational shortcuts for the unconstrained spline do not carry over. Minimization of $\mathrm{AdCV}(\lambda)$ is still feasible however, taking roughly 10 minutes $(n = 21)$ on a SUN SPARCstation 1.

One reason for considering this particular version of cross-validation was to compare the performance of constrained and unconstrained smoothing. The first and last values were omitted from the criterion because these were not available from our constrained spline software. We used ordinary cross-validation rather than generalized cross-validation because of the difficulties of identifying a linear approximation to the constrained spline solution. The generalized cross-validation function (Craven and Wahba 1979, Wahba 1990), for the unconstrained spline is

$$GCV(\lambda) = \frac{1}{n}\left[\sum_{k=1}^{n}\left(y_k - \hat{f}_\lambda(t_k)\right)^2\right] / \left[(1 - \text{trace}(A(\lambda))/n\right]^2$$

and is based partly on the property of the estimate being a linear (or approximately linear) function of $\underset{\sim}{y}$. Although Villalobos and Wahba suggest a substitute for trace($A(\lambda)$) in the constrained case, it is still is an approximation. Rice (1986) has suggested an alternative method for choosing lambda when estimating a derivative based on minimizing a nearly unbiased estimate of the integrated mean square error. In this case also, it was not clear how to extend this method to constrained spline estimates. Therefore, we chose to use "brute force" cross-validation because at least it was clear what the criterion was estimating.

2.4 Testing for Differences Among Rate Curves

In our application we are specifically interested in testing whether a set of rate curves generated from different depletion curves are the same. In this section we propose a simulation-based test for comparing rate curve estimates.

Suppose that one has M depletion experiments yielding rate curve estimates $\{\hat{\Phi}_1, \ldots, \hat{\Phi}_M\}$, corresponding residual sums of squares $\{RSS_1, \ldots, RSS_M\}$, and effective number of parameters $\{\nu_1, \ldots, \nu_M\}$ (from the trace of $A(\hat{\lambda})$). Let $\overline{\Phi}$ denote an "average" rate curve across these experiments. Although there are several ways to construct $\overline{\Phi}$, one reasonable way is to fit a cubic smoothing spline to the data set obtained by combining all pairs of estimated concentrations and rates across the the different experiments. In this manner the averaging process uses rate estimates only at observed concentrations. This is a crude way of weighting the depletion curves according to the number of observed concentrations in a particular range. Under the assumption that trace($A(\hat{\lambda})$) is a reasonable measure of the degrees of freedom associated with the model for a depletion curve, consider the pooled estimate of σ^2:

$$s_p^2 = \sum_{k=1}^{M} RSS_k / (N - \nu_T)$$

where $\nu_T = \sum_{k=1}^{M} \nu_k$ and $N = \sum_{k=1}^{M} n_k$. With these definitions let

$$F^* = \frac{\sum_{k=1}^{M} \int (\hat{\Phi}_k - \bar{\Phi})^2 dc}{s_p{}^2} . \tag{2.3}$$

The definition of F^* is analogous to the usual one-way ANOVA F statistic except that an L_2 measure replaces the simple difference of the treatment and grand means.

We wish to test the null hypothesis that the rate curves are equal against the alternative that they are not all equal. Clearly the distribution of F^* will not be of a simple known form even under this null hypothesis and is more easily approximated by simulation methods. These simulations are based on the assumption that $\bar{\Phi}$ is the true rate curve across all experiments and the added assumption that the measurement error is normally distributed with variance $\hat{\sigma}^2 = s_p{}^2$. Given these specifications for the model, pseudo data is generated for the individual depletion experiments and the rate curve estimates are computed exactly as they were with the original data. (This includes using GCV for selecting the smoothing parameter). For the set of M simulated depletion experiments the rate curve estimates are checked for positivity in the range of the test statistic. That is, if all of the depletion curve estimates are monotonic, then the test statistic at (2.3) is computed. Let F_B^* denote this (censored) random variable. (This condition is checked because it is assumed that the test statistic is only applied over a range where all the estimated rate curves are functions.) Percentage points from the distribution of F_B^* can be estimated and are used to set approximate critical values for the test statistic evaluated at the actual data.

It should be noted that the reference distribution for the test statistic is only an approximation to the true distribution since it is computed with respect to a pooled estimate of Φ rather than the true rate curve. Simulation methods using the estimated spline with normal errors have proved to be useful in estimating the distribution of the estimated smoothing parameter (Nychka 1991). This experience suggests that the simulated distribution of F_B^* may also be a valid approximation to the distribution of the test statistic. We are unable to give a theoretical justification at this time.

2.5 Consistency of the Rate Curve Estimate

In this section, a simple proof is outlined to establish the consistency of a nonparametric estimate of Φ. Although asymptotic properties of the estimator may seem irrelevant for this small sample application, the upper bound on the error of the estimate helps to identify features of the problem where the uptake rate is difficult to estimate.

First some notation is needed. Let $\| h \|_{\mathcal{A}} = \sup\{h(u): u \in \mathcal{A}\}$ and let $\mathcal{T}$ and $\mathcal{C}$ be intervals such that $f:\mathcal{T} \to \mathcal{C}$ and $f^{-1}:\mathcal{C} \to \mathcal{T}$ are bijective (this is guaranteed by the monotonicity of the depletion curve). We first state a general result and then make some specific comments related to the asymptotic properties of smoothing splines.

Theorem 1. Suppose that $\hat{f}$ is a monotonic estimate of f for the model at (1.1) with two continuous derivatives. Let $\hat{\Phi} = \hat{\Phi}(c) = \hat{f}'(\hat{f}^{-1}(c))$. Then

$$\| \hat{\Phi} - \Phi \|_{\mathcal{C}} \le \| \hat{f}'' \|_{\mathcal{T}} \; \| 1/\hat{f}' \|_{\mathcal{T}} \; \| f - \hat{f} \|_{\mathcal{T}} + \| \hat{f}' - f' \|_{\mathcal{T}}. \tag{2.4}$$

Proof. See Appendix A.1.

The convergence rate for estimates of the derivative will be slower than the rate for estimates of the function. Therefore the rate of convergence to zero of $\| \hat{\Phi} - \Phi \|_{\mathcal{C}}$ as $n \to \infty$ and $\lambda(n) \to 0$ will be dominated by $\| \hat{f}' - f' \|_{\mathcal{T}}$. Although the first term is asymptotically negligible, one might expect that this term is important when considering depletion curves. The tail of the depletion curve tends to have a small derivative for large values of t. If the curvature is constant then a derivative of f that is close to zero could inflate $\| 1/\hat{f}' \|_{\mathcal{T}}$ in this first term. An alternative estimator to adjust for flat tails of the depletion curve is suggested in Section 5.

For cubic smoothing splines Cox (1984) has given conditions on λ and f such that $\| \hat{f}'_{\lambda} - f' \|_{\mathcal{T}} \xrightarrow{P} 0$ as $n \to \infty$. Moreover, if f' is bounded away from zero then the unconstrained spline estimate will be monotonic with probability one as $n \to \infty$. For this reason enforcing monotonicty constraints should have no impact on the convergence rate of the estimate.

3. A SOLUTION DEPLETION EXPERIMENT

3.1 Background

An uptake experiment begins with an initial concentration, c_0, of the solution that surrounds the plant roots. The depletion period is kept as short as possible so that the capacity of the uptake system is not altered. Despite these controls, one question is whether the resulting uptake rate curve depends on the initial concentration. In the next two subsections we describe and analyze solution depletion data for nitrate absorption by the roots of corn seedlings (8-day old maize, P-3320) to study the effect of different initial concentrations on the shape of the uptake curve.

Suppose that Φ and Φ^* are the resulting rate curves for depletion experiments based on initial concentrations: $c_0 > c_0^*$. Under what circumstances will Φ and Φ^* coincide? If these curves were equal it would suggest that under these conditions the absorption mechanism does not depend on previous absorption of the nutrient. That is, when the root material has depleted the solution from c_0 to c_0^* the subsequent nutrient uptake is the same as a solution depletion experiment with an initial concentration of c_0^*. Although depletion experiments are relatively easy to carry out, the dependence of Φ on c_0 has some practical consequences that could limit their use. For example, one should not directly compare depletion results obtained at different initial concentrations if the underlying rate curves are different.

3.2 Experiment Design

Eight solution depletion experiments were carried out following a 4×2 factorial design. There were four initial concentrations, c_0 (50, 100, 200 and 500 μM), and two different nitrate exposure levels of the growing medium (low and high). The latter factor was included because it is known that the rate of nitrate uptake depends on the nitrate level in the growing medium. Each depletion was based on four plants in a container with 200ml of nutrient solution. The depletions were carried out simultaneously under constant light conditions over a 1-8 hour time period; the solutions were sampled at 21 different times. The details of the actual experiment were planned and performed by William Jackson and coworkers, Department of Soil Science, North Carolina State University.

Traditionally, the sampling times in a solution depletion experiment are chosen to be equally spaced. We designed the eight experiments to have unequally spaced sampling points with more samples taken at the onset of the experiment. Typically a depletion curve decreases rapidly at high concentrations and changes relatively slowly at low concentrations. The disadvantage of equally spaced sampling points is that fewer observations are taken when the concentration is high and changing rapidly than at lower concentrations where the depletion curve has less structure. Based on data from previous depletion experiments, we designed sequences of sampling points yielding measurements that were approximately equally spaced with respect to concentration rather than time. (The details of this procedure are in Meier 1990).

3.3 Analysis of the Data

For each of the eight data sets displayed in Figure 1, we examined four different estimates of Φ:

(1) Parametric estimate based on the Michaelis-Menten model

(2) Unconstrained cubic smoothing spline with λ determined by generalized cross-validation, GCV

(3) Unconstrained cubic smoothing spline with λ determined by adjusted generalized cross-validation, AdGCV

(4) Monotonic spline estimate defined at (2.2) with λ determined by adjusted cross-validation, AdCV.

Figures 2 and 3 display the rate curve estimates in a manner so that they can be easily compared. The plots in Figure 2 indicate that there are some differences between the Michaelis-Menten model for the rate curves and a nonparametric estimate. In particular, the rate curves for the highest levels of initial concentration appear to have much more structure than the sharp step suggested by the Michaelis-Menten estimate. It is also worth noting that for each of the eight treatments, the Michaelis-Menten estimate is consistently higher than the nonparametric estimate at lower concentrations and consistently lower than the nonparametric estimate at the higher concentrations. Three of the eight unconstrained depletion curve estimates were monotone functions. In fact, except for the lowest concentrations ($< 2.5\ \mu$M), all of the unconstrained depletion curve estimates were monotone. The estimates based on the adjusted cross-validation criterion (Figure 3) yield rougher estimates of the rate curve than those based on GCV (Figure 2). The sharp upward spike at the right boundary of some of the monotone estimates is most likely due to boundary effects and the loops in the unconstrained estimate at low concentrations are due to lack monotonicity. Besides these differences at the ends of the range, the two estimates have a similar appearance.

One possible explanation for the bumps in the nonparametric estimates is that it is spurious structure due to undersmoothing. That is, for this experimental data the lambda that minimizes the AdGCV function and the lambda that minimizes the AdCV function underestimate the appropriate amount of smoothness necessary to yield a smooth rate curve. An alternative method for choosing lambda is to choose that value of lambda so that the variance estimated from the smoothing spline estimate, $\hat{\sigma}_s^2$ (see Craven and Wahba 1979), is equal to or at least as large as an independent (experimental) estimate of σ^2 (Reinsch, 1967). For this experiment, the independent estimate of σ^2 was 0.04. All of the rate estimates except for the case c_0=200, LOW nitrate exposure, yielded estimates where $\hat{\sigma}_s^2$ was slightly larger than 0.04. For the one case where $\hat{\sigma}_s^2$ was less than 0.04, a new rate curve estimate was computed where lambda was chosen so that $\hat{\sigma}_s^2 \approx 0.04$. The resulting new estimate could not be distinguished from the original estimate. Thus the estimated rate curves appear to be in agreement with the expected amount of measurement error.

From the simulation study reported in Section 4 it was found that a monotonic estimate was less accurate (larger average squared error) than an unconstrained estimate. Also, given the fact that the unconstrained estimates were actually monotonic over a wide range of concentrations the treatments were compared using the unconstrained estimates with λ determined by generalized cross-validation.

Figure 4 plots the estimated rate curves separated into the low and high nitrate pretreatment groups. It is expected that the pretreatment will have some influence on the uptake rates. However, it is also of interest to determine whether there is a significant difference among the rate curves within each pretreatment group. A difference would suggest some dependence on the initial concentration. As the simplest hypothesis the rate curves were compared over a restricted range of concentrations where the lower bound was set to insure the estimates were functions (2.5μM) and the upper endpoint was the highest concentration common to all the experiments (50 μM) (Figure 5). We used the method described in Section 2.3 to test for differences among these curves and the results are reported in Table 1.

For each test the distribution of the test statistic under the null hypothesis was calculated using 10,000 simulated data sets. Due to the condition that the depletion curve be monotonic over the range (2.5 - 50 μM) the actual number of samples used in calculating F^*_B varied between 5462 and 8770. Perhaps the most important test result is the significant differences among the rate curves for the high pretreatment group of depletions.

In both pretreatment groups we noticed that the rate curves tend to be ordered with the higher initial concentrations yielding slightly higher rates (at the same solution concentration). The only departure in this pattern is the 100 ml depletion curve associated with the high nitrate pretreatment group. This group of four curves had a significantly small p-value when testing for equality but we were concerned about the dependence of this result on the high curve for the 100ml depletion. A test was done with this rate curve excluded and the results are also reported. We found that the p-value for this reduced set of curves increases roughly by a factor of four to .033.

4. SIMULATION EXPERIMENT

4.1 Objectives and Design

We conducted a simulation study to examine two aspects of estimating rate equations using the smoothing spline estimator, $\hat{\Phi}_\lambda$. We were interested in the value of using cross-validation for determining λ, and in the possible improvement in average squared error and smoothness of the rate curve when monotonicity constraints are imposed on the estimated depletion curve.

This study consists of 12 cases following a 4×3 factorial design consisting of four test functions (listed in Appendix A.2) and three levels of σ (0.02, 0.2, 0.5). The particular test functions (rate curves) used were chosen to be realistic for maize nitrate uptake and the middle level of σ, σ=0.20, is one that is typically observed in the laboratory for solution depletion experiments. For each of these 12 test cases, 240 samples of size 20 were simulated according to the model (1.1) with normal errors. For each of these samples the estimates 2 and 4 listed in Section 3.2 were calculated along with two other "optimal" estimates described below.

A global criterion to measure the closeness of the rate estimates to Φ is the integrated mean squared error,

$$\mathrm{IMSE}(\hat{\Phi}) = \int \left(\hat{\Phi}(u) - \Phi(u)\right)^2 du.$$

As an approximation to this integral we computed the average squared error (ASE) to assess smoothness of $\hat{\Phi}$, where

$$\mathrm{ASE}(\hat{\Phi}) = \frac{1}{m} \sum_{k=1}^{m} \left(\hat{\Phi}(u_k) - \Phi(u_k)\right)^2,$$

$u_k = \hat{f}(t_k)$ and m=500. This quantity is an approximation to the integrated mean squared error divided by $(u_m - u_0)$. The grid points $\{t_k\}$, $1 \leq k \leq 500$ and were chosen to yield equally spaced values of the *true* depletion curve, $\{f(t_k\}, 1 \leq k \leq 500\}$. In this manner $\hat{\Phi}(\hat{f}(t_k))$ is uniquely defined for all t_k and will be evaluated on a roughly equally spaced grid with respect to estimated concentration. The ASE was computed for both constrained and unconstrained rate estimates and is the basis for comparison among the estimators.

Finally, in order to calibrate the data-based methods of determining the smoothing parameter, the optimal choice for λ (with respect to ASE) was also calculated. These minimizations were carried out by a coarse grid search that was refined using a golden section minimization algorithm.

4.2 Results of Simulation Study

Figure 6 and Table 2 provide a summary of the simulation results that are pertinent to the analysis of the uptake experiment. The figure gives an overview of the accuracy of the rate curve estimates under the 12 cases and across four different estimates.

For data-based methods of estimating λ (labeled as gcv u and cv m) the distribution of the ASE's have roughly the same variability. One systematic difference, however, is that the median ASE associated with the data-based unconstrained estimate is smaller than the ASE for the constrained estimate. This pattern is also true for the rate curve estimates using the optimal value for λ. Here we see similar scatter but the distribution of the ASE for unconstrained estimate tends to be shifted to the left (or to lower values) than that of the constrained estimate. Another important feature of these results is the increasing variability and skewness of the ASE for the data-based estimates when σ increases. Usually these poor estimates are associated with values for λ that substantially undersmooth the data.

Table 2 lists the median values for these results and quantifies the general patterns apparent in the figure. For example, when $\sigma = .2$ estimating λ by GCV for the unconstrained estimate has the effect of inflating the square root of the median ASE in a range of 20% to 60%.

5. DISCUSSION AND CONCLUSIONS

Our analysis of a depletion experiment differs from most other researchers in that we do not assume the rate curve must satisfy a particular parametric form. Using a nonparametric regression estimate of the rate curve, we suggest an inference procedure based on Monte Carlo simulation. The nonparametric estimates of the rate curves appear to differ qualitatively from a Michaelis-Menten model. While we did not specifically test for inadequacy of the Michaelis-Menten model, the results for the depletions with largest initial concentrations (see $C_0 = 500$ μM in Figure 2) clearly suggest a different shape for the rate curve based on a nonparametric analysis. Based on the tests for equality among rate curves there is strong statistical evidence that the rate curves differ among the group with high exposure to nitrate. We were not able to find similar differences among the low exposure group.

Several qualifications need to be made to such conclusions. One explanation of the observed differences is that over longer periods of depletion time the capacity of the uptake system may actually be changing. Thus in these experiments the effect due to initial concentration is confounded with the length of the depletion experiment (see Figure 1). Another issûe is the lack of replication for this experiment. Systematic effects among containers cannot be estimated from the single depletion experiment done at each of the four initial concentrations. Another problem with these conclusions is the influence that the depletion with $c_0 = 100$ ml has on the over significance of the hypothesis test. Although there are possible scientific explanations for the unexpected ordering of the 100 ml rate curve relative to the others, it is also possible that this outcome was due to uncontrolled variables in the depletion experiment. For this reason the p-value associated with the high exposure group omitting this depletion is a more conservative measure of the significance. Note that in the low exposure group there is still some evidence for differences among the rates and when interpreting these test results, one should keep in mind the low power associated with an overall test.

There are some other aspects of these data that might be explored. For example, one might test for departures from a Michaelis-Menten rate model or test for pairwise differences among rate curves. Although the simulation method for obtaining the test statistic distribution is computationally intensive, it gives more flexibility in making comparisons. For example when comparing rate curves pairwise it is easy to adjust the level of significance due to multiple comparisons. Besides further statistical analysis our results suggest some additional experiments, most notably investigating the $c_0 = 100\mu M$ case for the high pretreatment group. In fact the original experiment was intended as the first half of a larger study with the intent that preliminary results would guide the design of the second half.

One interesting result from the simulation study was the decreased accuracy associated with constrained estimates. Although one might argue that enforcing monotonicity adds more information to the estimate, this was not the case for the sample sizes and rate curves that are appropriate for depletion experiments. In addition, we found that the unconstrained estimates were monotonic over the most relevant range of concentrations. Thus the need to enforce constraints was not necessary.

Cross-validation yields smoothing parameter estimates that are reasonable for most cases. For ordinary smoothing splines it has been found that the square root of the expected average squared error, when treated like a standard error will give a reasonable pointwise confidence interval for the curve estimate (Nychka 1988). Stretching this property to the rate curve setting, one might view the square root of the ASE as a relative measure of the width of pointwise confidence intervals for Φ. Under this correspondence, estimating λ by GCV inflates the confidence interval width by at most 60% when $\sigma = .2$. This modest loss in efficiency may be acceptable given the difficulty of using other methods of smoothing parameter selection.

The simulation results also point to substantial variability in the individual estimates of λ. Evidence of this problem is the the skewed distributions of ASE for the large values of σ in Figure 6. Recently some alternatives to cross-validation have been proposed that may have much less variability (Hall, Park and Marron 1990). These procedures, however, require detailed calculations of the asymptotic bias and variance of the curve estimate. Such formula are difficult to derive for splines and may not be accurate due to the differences between cubic smoothing splines and second order kernel estimators.

One basic objection to the formulation of the spline estimate is that the roughness penalty is associated with the depletion curve rather than the function of interest, Φ. It is possible to formulate a variational problem that depends directly on the rate function. Let K be an operator such that $f=K(\Phi)$ will be the solution to the rate equation with $f(0) = c_0$. Under the assumption that c_0 is known one might consider minimizing

$$\mathcal{L}(\Phi) = \frac{1}{n} \sum_{i=1}^{n} (y_i - K(\Phi)(t_i))^2 + \lambda \int_{[0,c_0]} (\Phi''(u))^2 du \tag{5.1}$$

over all rate curves such that $\Phi \geq 0$, $\Phi(0) = 0$ and $\int_{[0,c_0]} (\Phi''(u))^2 du < \infty$.

In this way the roughness constraint is placed directly on the rate curve. Part of the problem with monotonicity arises for low concentrations where the depletion curve is essentially following an exponential decay rate. One advantage of this formulation is that in the case of $\lambda = \infty$, Φ will be a linear function. If Φ is linear then the corresponding depletion curve will be an exponential function and may provide a better representation of the data for low concentrations.

From a theoretical point of view this direct estimate of Φ may seem better suited for this problem. We did not pursue this method, however, due to numerical considerations. Besides the basic issue of whether a minimizer of (5.1) even exists, computation of the estimate would be formidable. Evaluating K at a particular Φ would require the numerical solution of the rate equation for f. In addition, it would be difficult to estimate the smoothing parameter since this is a nonlinear problem. Given the rapid growth in computing power, however, we hope that such problems will soon become feasible. Differential equations describing the rate of a process are a useful tool for understanding biological systems, and we believe that this is rich field for the application of nonparametric (and computer intensive) methods.

APPENDIX

A.1 Proof of Theorem 1

First note that from the elementary properties of supremum norm:

$$\| f \cdot g \|_{\mathcal{A}} \leq \| f \|_{\mathcal{A}} \| g \|_{\mathcal{A}} \quad \text{(P1)}$$

and if f and f^{-1} are bijective, then

$$\| f^{-1} \|_{\mathcal{C}} = \| f^{-1} \circ f \|_{\mathcal{T}} . \quad \text{(P2)}$$

Define $\hat{g} \equiv \hat{f}^{-1}$. By the triangle inequality and then by (P2)

$$\begin{aligned} \| \hat{\Phi} - \Phi \|_{\mathcal{C}} &= \| \hat{f}' \circ \hat{g} - f' \circ f^{-1} \|_{\mathcal{C}} \\ &\leq \| \hat{f}' \circ \hat{g} - \hat{f}' \circ f^{-1} \|_{\mathcal{C}} + \| \hat{f}' \circ f^{-1} - f' \circ f^{-1} \|_{\mathcal{C}} \\ &= \| \hat{f}' \circ \hat{g} \circ f - \hat{f}' \circ f^{-1} \circ f \|_{\mathcal{T}} + \| \hat{f}' - f' \|_{\mathcal{T}} . \end{aligned}$$

Applying the Mean Value Theorem to $\hat{f}'$ and then applying (P1), the first term on the right hand side becomes

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

Substituting $I = \hat{f}^{-1} \circ \hat{f} \equiv \hat{g} \circ \hat{f}$, applying the Mean Value Theorem to $\hat{g}$ and again applying (P1), this last expression becomes

$$\begin{aligned} \| \hat{f}' \|_{\mathcal{T}} \, \| \hat{g} \circ f - \hat{g} \circ \hat{f} \|_{\mathcal{T}} &\leq \| \hat{f}' \|_{\mathcal{T}} \, \| \hat{g}' \|_{\mathcal{T}} \, \| f - \hat{f} \|_{\mathcal{T}} \\ &= \| \hat{f}' \|_{\mathcal{T}} \, \| 1/\hat{f}' \|_{\mathcal{T}} \, \| f - \hat{f} \|_{\mathcal{T}} \end{aligned}$$

where $\hat{g}' \equiv \frac{d}{dc} \hat{f}^{-1}(c) = 1/\hat{f}'(t)$ with $c=f(t)$. Combining all of the pieces,

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

A.2 Rate Curves Used in Simulation Study

Observations were generated from the depletion curves corresponding to the four rate functions below:

$\Phi(u) = \frac{12 \cdot u}{40+u}$, $c_0=50$ (Michaelis-Menten (LOW))

$\Phi(u) = \frac{12 \cdot u}{40+u}$, $c_0=500$ (Michaelis-Menten (HIGH))

$\Phi(u) = -0.076 \cdot u$, $c_0=50$ (Linear (LOW))

$\Phi(u) = -0.017 \cdot u$, $c_0=500$ (Linear (HIGH)).

To obtain the corresponding depletion curve, for each of these rate curves, the first order system $\Phi(f) = -f'$, $f(0) = c_0$ was solved for f. The Michaelis-Menten equations do not have an explicit, closed form solution and therefore were solved numerically using fourth order Runge-Kutta methods (Boyce and DiPrima 1969). The numerical solution consisted of the pairs $(x_h, f(x_h))$ where $c_0=x_0 \leq x_1 \leq \ldots \leq x_n$, and the step size, $x_{i+1}-x_i$ was chosen to be less than or equal to 0.01. The solution was evaluated up to x_n where $f(x_n)$ was less than or equal to 2 for the first time. This corresponds to a solution depletion experiment where the solution is sampled until its concentration falls below 2 uM. To obtain 20 time points, $\{t_k, 1 \leq k \leq 20\}$ that yield equally spaced $\{f(t_k\}, 1 \leq k \leq 20\}$, the following method was used. An interpolating spline was fit to the "inverse" data $\{(f(x_h), x_h)\}$, and then this interpolating spline was evaluated at 20 equally spaced points, $\{f(t_k), 1 \leq k \leq 20\}$. The result was a grid of unequally spaced time points, $\{t_k, 1 \leq k \leq 20\}$.

Unlike the Michaelis-Menten equations, the linear equations can be solved explicitly as $f(t) = 50\,e^{-0.076\,t}$ (LOW) and $f(t) = 100\,e^{-0.017\,t}$ (HIGH). These particular exponential functions were chosen because they minimized the residual sum of squares between the function and the numerical solution to the corresponding Michaelis-Menten rate equation. The 20 time points that yield equally spaced $\{f(t_k), 1 \leq k \leq 20\}$ were found directly from the exponential form.

Table 1 Results of testing for equality of rate curves obtained from different initial concentrations for a particular exposure level.

	Null Hypothesis			
	Low	High	All	High excluding $c_0 = 100\ \mu M$
P-value[1]	0.108	0.008	0.017	0.033
%monotone[2]	87.7	54.6	62.2	61.3

[1] Approximate p-value based on the simulated distribution of F^*.

[2] The percent of simulated samples out of 10,000 that yielded monotone estimates for each treatment.

Table 2 Simulation results comparing several nonparametric estimates of the rate curve.

Test Function[1]	Value for λ	Median average squared error[4] unconstrained	monotone	unconstrained /monotone
M-M (LOW)				
$\sigma = 0.02$	CV[2]	.632e-4	.683e-4	.960
	Opt.[3]	.311e-4	.406e-4	.822
$\sigma = 0.20$	CV	.904e-3	.955e-3	.992
	Opt.	.368e-3	.396e-3	.958
$\sigma = 0.50$	CV	.257e-2	.292e-2	.993
	Opt.	.129e-2	.131e-2	.982
Linear (LOW)				
$\sigma = 0.02$	CV	.760e-5	.127e-4	.575
	Opt.	.579e-5	.114e-4	.516
$\sigma = 0.20$	CV	.193e-3	.263e-3	.757
	Opt.	.132e-3	.177e-3	.790
$\sigma = 0.50$	CV	.673e-3	.868e-3	.868
	Opt.	.460e-3	.527e-3	.885
M-M (HIGH)				
$\sigma = 0.02$	CV	.209e-3	.302e-3	.717
	Opt.	.135e-3	.214e-3	.686
$\sigma = 0.20$	CV	.346e-2	.369e-2	.993
	Opt.	.136e-2	.155e-2	.909
$\sigma = 0.50$	CV	.836e-2	.911e-2	.100
	Opt.	.344e-2	.?51e-2	.978
Linear (HIGH)				
$\sigma = 0.02$	CV	.?03e-3	.573e-3	.365
	Opt.	.151e-3	.698e-4	2.160
$\sigma = 0.20$	CV	.102e-2	.147e-2	.714
	Opt.	.689e-3	.110e-2	.666
$\sigma = 0.50$	CV	.383e-2	.426e-2	.918
	Opt.	.175e-2	.240e-2	.769

[1] Test functions are listed in Appendix 2.

[2] Smoothing parameter found by either GCV (unconstrained spline) or AdCV (monotone spline)

[3] Optimal choice of smoothing parameter found by minimizing the average squared error (ASE).

[4] Median values of ASE's and ratios of ASE's based on 240 simulation estimates per test function.

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.

de Boor, C. (1978), *A Practical Guide to Splines*, New York: Springer-Verlag.

Boyce, W. E. and DiPrima, R. C. (1969), *Elementary Differential Equations and Boundary Value Problems*, Second edition, New York: Wiley.

Brady, N. C. (1984), *The Nature and Properties of Soils*, Ninth edition, New York: Macmillan Publishing Company.

Clarkson, D. T. (1986), "Regulation of the Absorption and Release of Nitrate by Plant Cells: A Review of Current Ideas and Methodology," *In* H. Lambers, J.J. Neeteson and I. Stulen, eds., Fundamental, Ecological and Agricultural Aspects of Nitrogen Metabolism in Higher Plants. Martinus Nijhoff Publishers, Dordrecht, 3-27.

Claassen, N. and Barber, S. A. (1974), "A Method for Characterizing the Relation Between Nutrient Concentration and Flux Into Roots of Intact Plants," *Plant Physiology* 54, 564-568.

Cox, D. D. (1984), "Multivariate Smoothing Spline Functions" *SIAM Journal of Numerical Analysis* 21, 789-813.

Craven, P. and Wahba, G. (1979), "Smoothing Noisy Data with Spline Functions: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross-validation," *Numerische Mathematik* 31, 377-403.

Cullum, J. (1971), "Numerical Differentiation and Regularization," *SIAM Journal of Numerical Analysis* 8, 254-265.

Du Chateau, P. C., Nofziger, D. L., Ahuja, L. R. and Swartzendruber, D. (1972), "Experimental curves and rates of change from piecewise parabolic fits," *Agronomy Journal* 64, 538-542.

Epstein, E. (1972). *Mineral Nutrition of Plants: Principles and Perspectives.* New York: Wiley.

Epstein, E. and Hagen, C.E. (1952), "A Kinetic Study of the Absorption of Alkali Cations by Barley Roots," *Plant Physiology* 27, 457-474.

Eubank, R. L. (1988), *Spline Smoothing and Nonparametric Regression*, New York: Marcel Dekker, Inc.

Gasser, T. and Müller, H. G. (1984), "Estimating Regression Functions and Their Derivatives by the Kernel Method," *Scandinavian Journal of Statistics* 11, 171-185.

Hall, P., Park, B. U. and Marron, J. S. (1990), "Smoothed Cross-Validation," unpublished manuscript.

Jackson, W. A., Pan, W. L., Moll, R. H., Kamprath, E. J. (1986), "Uptake, Translocation and Reduction of Nitrate," *In* C.A. Neyra, ed, Biochemical Basis of Plant Breeding, Vol. II. Nitrogen

Metabolism. CRC Press, Boca Raton, FL, 73-108.

Jackson, W. A., Volk, R. J., Morgan, M. A., Pan, W. L. and Teyker, R.H. (1986), "Nitrogen Uptake and Partitioning by Roots," *In* J.C. Shannon, D.P. Knievel and C.D. Boyer, eds, Regulation of Carbon and Nitrogen Reduction and Utilization in Maize. The American Society of Plant Physiologists, 83-104.

Jackson, W. (1989), Personal communication.

Jacquez, J. A. (1985), *Compartmental Analysis in Biology and Medicine*, Ann Arbor: The University of Michigan Press.

Kelly, C. and Rice, J. (1990), "Monotone Smoothing With Application to Dose Response Curves and the Assessment of Synergism," *Biometrics* 46, 1071-1086.

Lawson, C. L. and Hanson, R. J. (1974), *Solving Least Squares Problems*, Englewood Cliffs: Prentice Hall.

Meier, K. (1990), "Estimating Rate Equations Using Nonparametric Methods," Ph.D. dissertation, Department of Statistics, North Carolina State University.

Michaelis, L. and Menten, M. L. (1913), "Die Kinetik der Invertinwirkung," *Biochem. Z.* 49, 333-369.

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.

Nychka, D. (1991), "Choosing a Range for the Amount of Smoothing in Nonparametric Regression," *Journal of the American Statistical Association*, to appear.

Ramsay, J. O. (1988), "Monotone Regression Splines in Action," *Statistical Science* 3, 425-441.

Reinsch, C. (1967), "Smoothing by Spline Functions," *Numerische Mathematik* 10, 177-183.

Rice, J. (1986), "Bandwidth Choice for Differentiation," *Journal of Multivariate Analysis* 19, 251-264.

Rice, J. and Rosenblatt, M. (1983), "Smoothing Splines: Regression, Derivatives and Deconvolution," *the Annals of Statistics* 11, 141-156.

Silverman, B. W. (1984), "Spline Smoothing: The Equivalent Variable Kernel Method," *The Annals of Statistics* 12, 898-916.

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.

Utreras, F. (1985), "Smoothing Noisy Data Under Monotonicity Constraints: Existence, Characterization and Convergence Rates," *Numerische Mathematik* 47, 611-625.

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. (1975), "Smoothing Noisy Data with Spline Functions," *Numerische Mathematik* 24, 383-393.

Wahba, G. (1990), *Spline Models for Observational Data*, SIAM, Philadelphia, PA.

Figure 1 Nitrate uptake by maize roots over time. The observed amount of nitrate (μmol) remaining in ambient solution is plotted as a function of time. The initial volume of each solution was 200 ml. The nitrate concentration of the ambient solution is in μM (1 μM = 1 μmol/1000 ml). LOW NO3 or HIGH NO3 indicates whether the plants were grown in a solution with low or high nitrate concentration prior to the experiment.

Figure 2 Comparison of nonparametric rate curve estimates (solid) and best fit of the Michaelis-Menten model (dashed). Each nonparametric rate curve estimate is based on an unconstrained estimate of the depletion curve; lambda is the minimizer of GCV(λ). The best Michaelis-Menten model is the minimum least squares fit of the numerically integrated Michaelis-Menten model. The uptake rate is given in units of μmol nitrate per gram root mass per hour.

Figure 3 Comparison of nonparametric rate curve estimates based on unconstrained (solid) and monotone (dashed) estimates of the depletion curve. For the unconstrained estimates, lambda is the minimizer of AdGCV(λ); for the monotone estimates lambda is the minimizer of AdCV(λ).

Figure 4 Dependence of rate curves on initial concentration. Each nonparametric rate curve estimate is based an unconstrained estimate of the depletion curve; lambda is the minimizer of GCV(λ). C_0 is the initial nitrate concentration of the ambient solution in μM.

Figure 5 A magnification of Figure 4. The vertical lines indicate the range where the test statistic at (2.3) was computed.

Figure 6 Simulation results comparing different nonparametric estimates of rate curves. Boxplots of the log average squared error (ASE) of 240 rate curve estimates for each of 4 different types of estimates: unconstrained estimate, lambda is the minimizer of AdGCV (gcv u); unconstrained estimate, lambda is the minimizer of ASE (ase u); monotone estimate, lambda is the minimizer of AdCV (cv u); and unconstrained estimate, lambda is the minimizer of ASE (ase m).

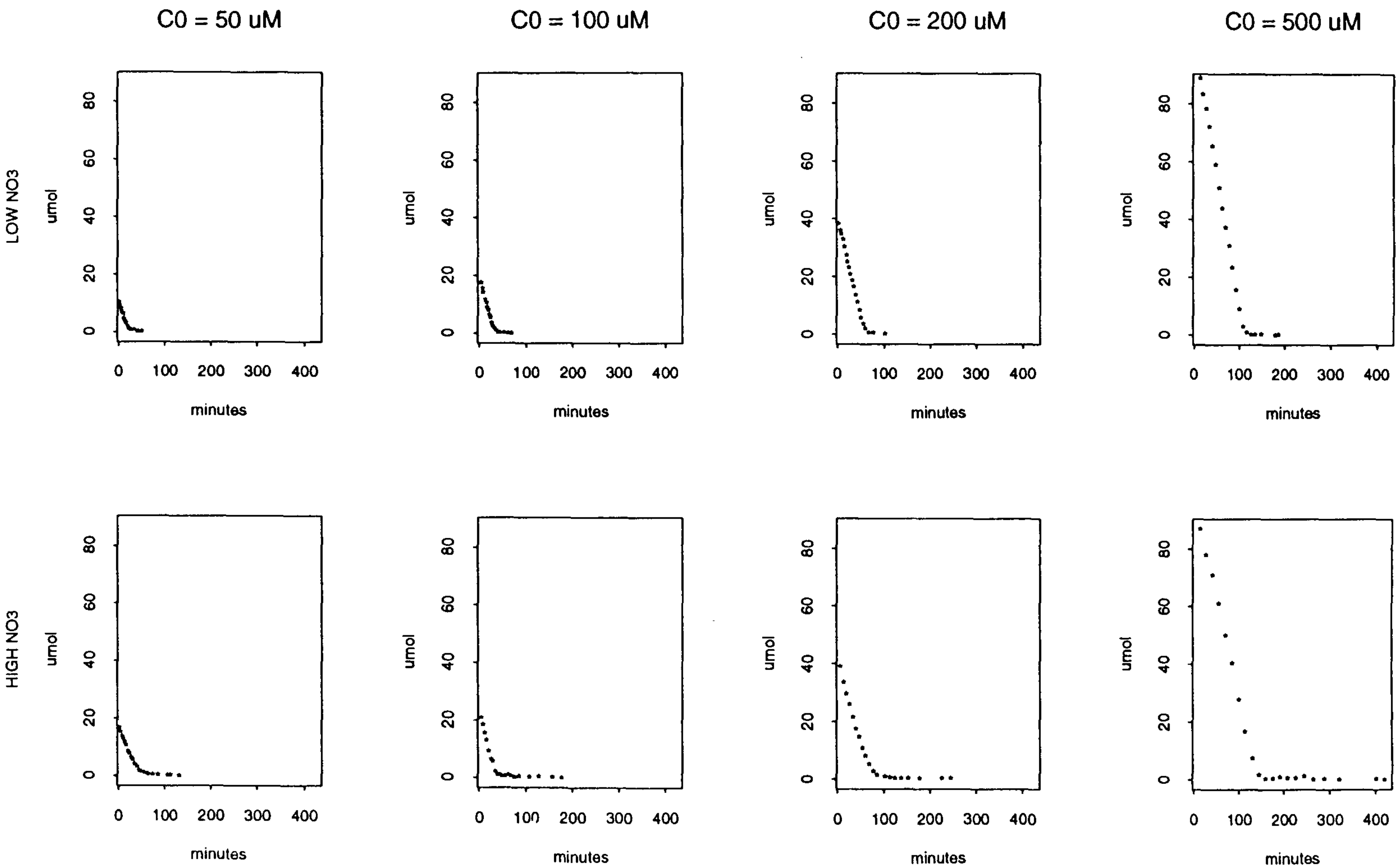

C0 = 50 uM
C0 = 100 uM
C0 = 200 uM
C0 = 500 uM
LOW NO3
HIGH NO3
umol
0 20 40 60 80
0 100 200 300 400
minutes

Figure 1

Figure 2

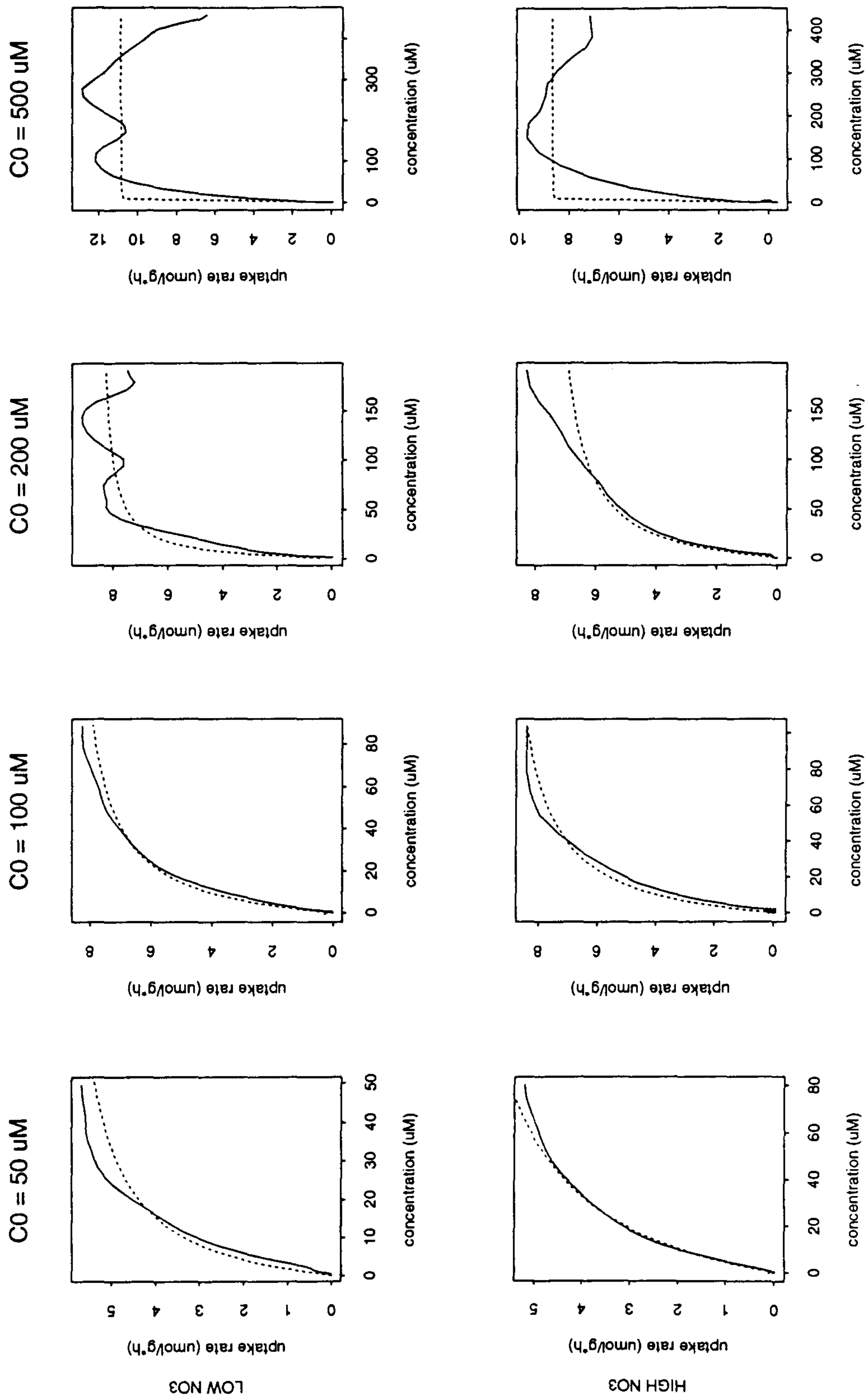
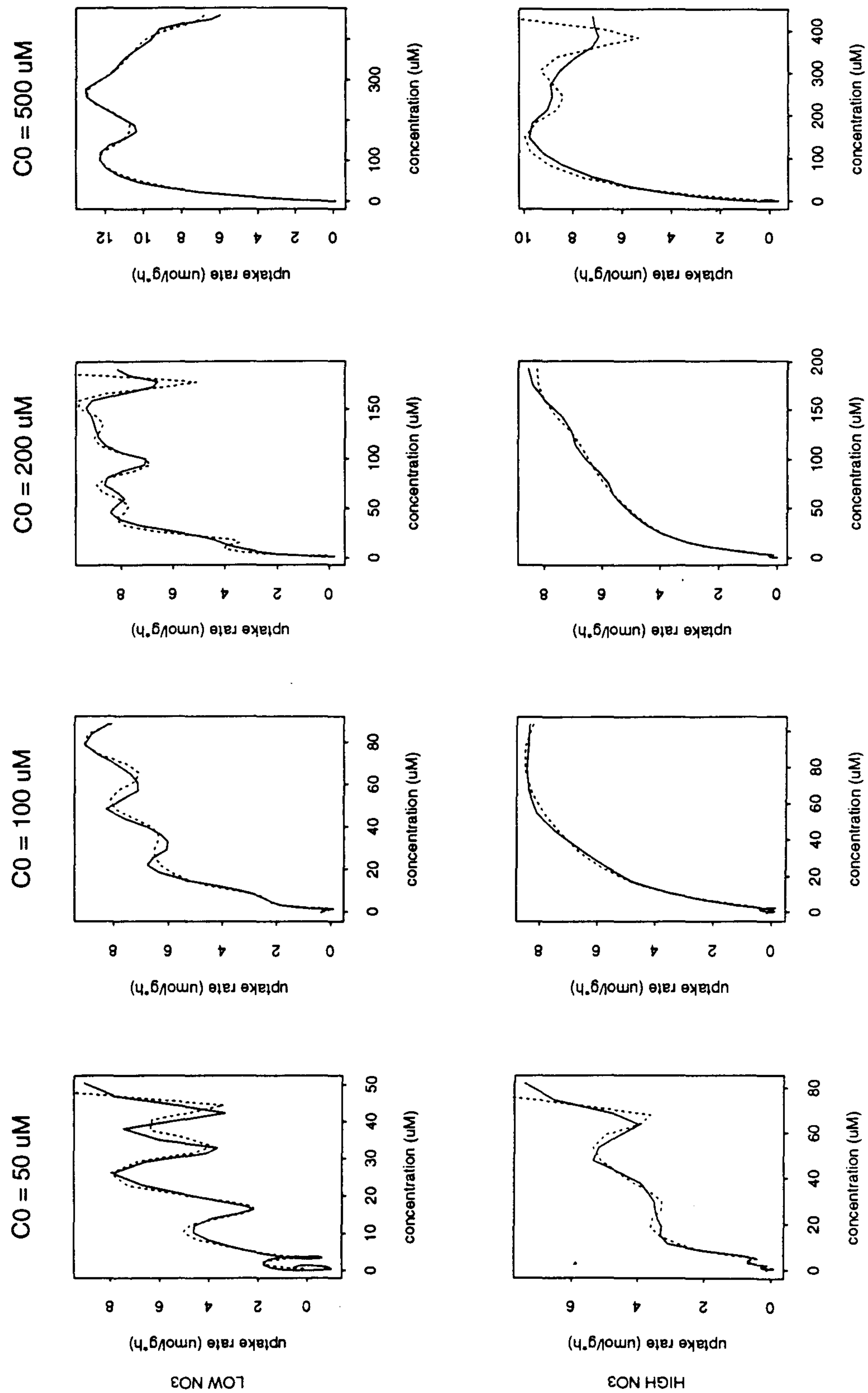

C0 = 500 uM
C0 = 200 uM
C0 = 100 uM
C0 = 50 uM
uptake rate (umol/g*h)
concentration (uM)
LOW NO3
HIGH NO3

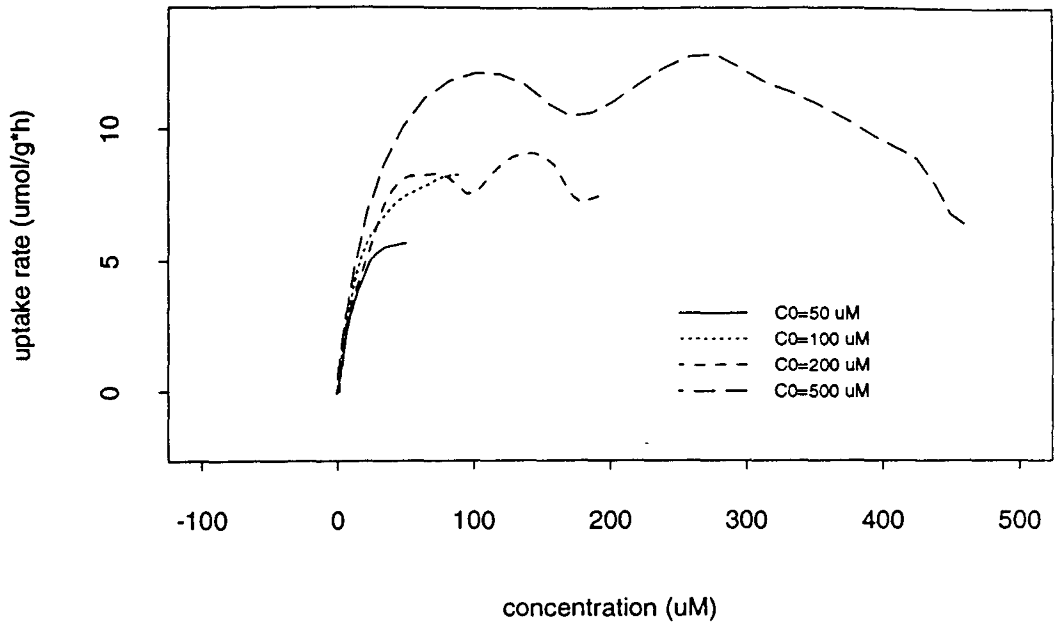
Figure 3



C0 = 50 uM
C0 = 100 uM
C0 = 200 uM
C0 = 500 uM
LOW NO3
HIGH NO3
uptake rate (umol/g*h)
concentration (uM)

Figure 4

LOW Nitrogen Exposure



HIGH Nitrogen Exposure

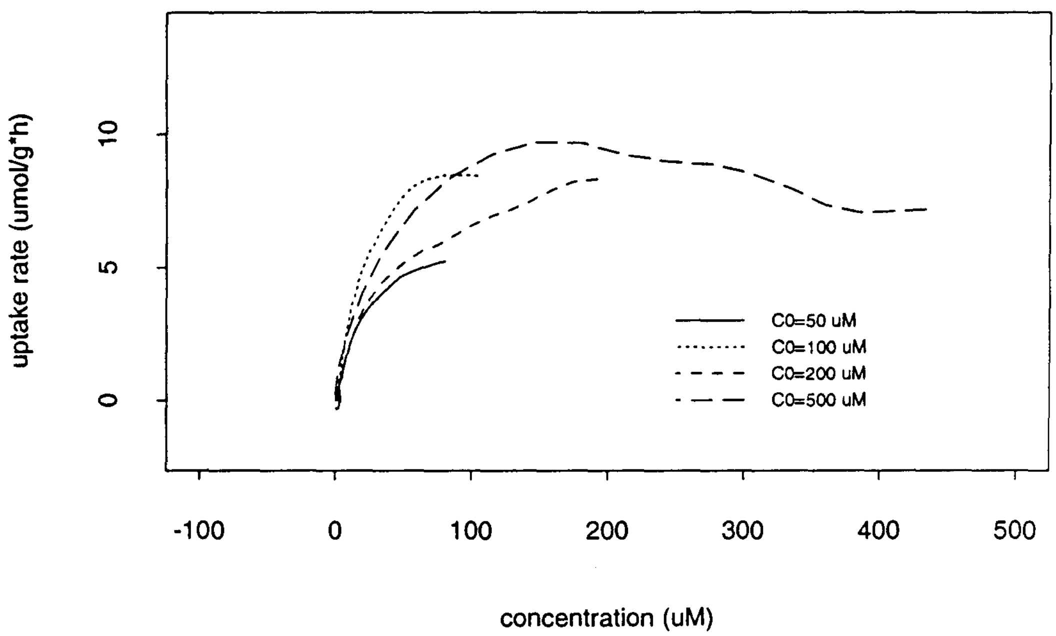
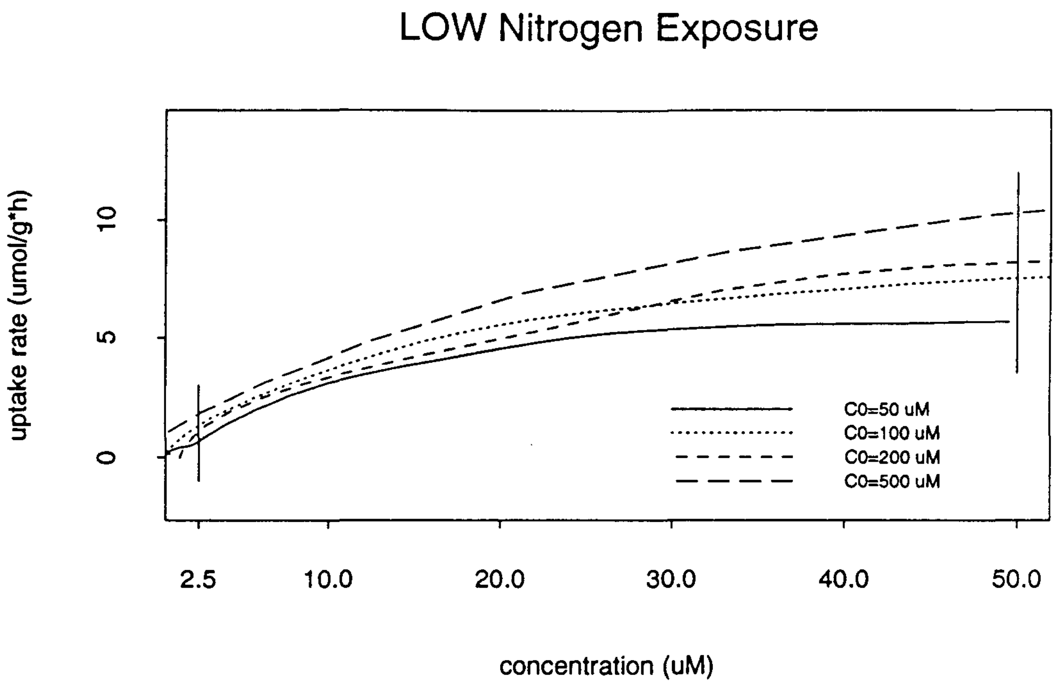
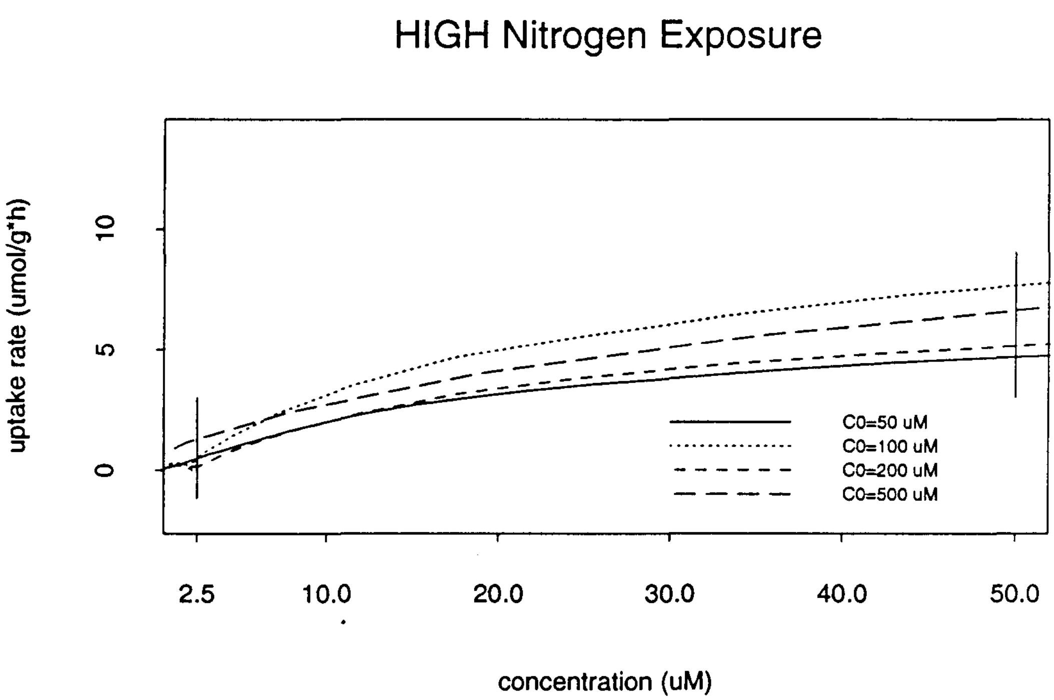


Figure 5


LOW Nitrogen Exposure
uptake rate (umol/g*h)
0
5
10
C0=50 uM
C0=100 uM
C0=200 uM
C0=500 uM
2.5
10.0
20.0
30.0
40.0
50.0
concentration (uM)


HIGH Nitrogen Exposure
uptake rate (umol/g*h)
0
5
10
C0=50 uM
C0=100 uM
C0=200 uM
C0=500 uM
2.5
10.0
20.0
30.0
40.0
50.0
concentration (uM)

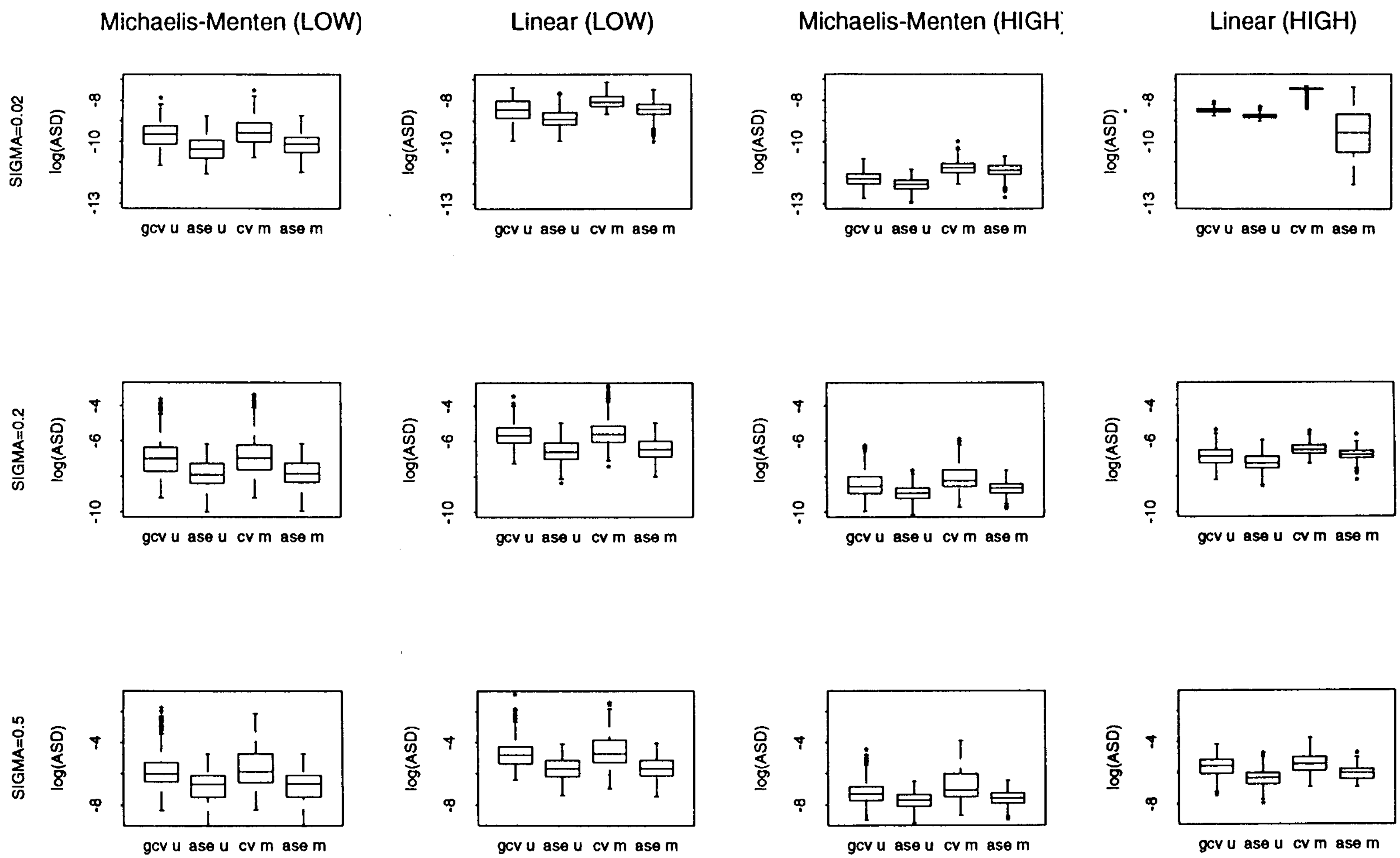

Michaelis-Menten (LOW)
Linear (LOW)
Michaelis-Menten (HIGH)
Linear (HIGH)
SIGMA=0.02
SIGMA=0.2
SIGMA=0.5
log(ASD)
gcv u
ase u
cv m
ase m

Figure 6