J. ENVIRON. SYS., Vol. 4(1), Spring 1974

NONLINEAR LEAST SQUARES TECHNIQUES FOR SYSTEM IDENTIFICATION IN WATER QUALITY

J.-G. BELIVEAU
Research Associate
Department of Civil Engineering
Columbia University

G. MATTINGLY
Assistant Professor
Department of Civil and Geological Engineering
Princeton University

ABSTRACT

In natural surface waters such as rivers and lakes, the supply of dissolved oxygen (DO) and the oxygen demand (BOD) are measurable quantities which determine the water's quality. Using specific water quality modeling systems and records of these measurable quantities, the important parameters governing the system response can be found. Once these parameters are determined, meaningful sets of controls may be imposed to keep water quality at or above acceptable standards. Many models have been proposed to represent the experimental observations. Most of these are variations of the classical Streeter-Phelps equation for the oxygen-sag relationship in rivers. The model which is considered in the present effort is due to Camp, and considers such effects (and the respective parameters) as sedimentation, (k_3); photosynthesis (A); runoff (R); reaeration rate (k_2); and the deoxygenation rate (k_1). The method of nonlinear least squares combined with eigenvalue perturbations and parametric differentiation is used for parameter estimation for cases with both BOD and DO data and for DO data only. Both numerically generated test cases and actual laboratory experiments are considered in this "inverse" procedure.

In natural surface waters such as rivers and lakes, the supply of dissolved oxygen (DO) and the biochemical oxygen demand (BOD) are measurable quantities which determine the water's quality. Using specific water quality modeling systems and records of these measurable quantities, the important parameters

doi: 10.2190/JCHN-0KRK-4JQN-EDMW
http://baywood.com

governing the system response can be found. Once these parameters are determined, meaningful sets of controls may be imposed to keep water quality at or above acceptable standards.

Many models have been proposed to represent the experimental observations [1]. Most of these are variations of the classical Streeter-Phelps equation for the oxygen-sag relationship in rivers. More complicated models of stream water quality would consider the system of partial differential equations resulting from the complete mass, momentum, and energy transport considerations [2]. Various assumptions conveniently reduce these equations to simple relationships which seem to model the behavior in rivers adequately. The model which we consider in the present effort is that due to Camp [3]. It considers such effects (and the respective parameters) as sedimentation, (k_3); photosynthesis, (A); runoff, (R); reaeration, (k_2); and deoxygenation, (k_1). These equations are written

$$\begin{aligned} \frac{dB}{dt} + (k_1 + k_3)B &= R \\ \frac{dD}{dt} + k_2 D &= k_1 B - A \end{aligned} \tag{1}$$

where B is the biochemical oxygen demand in parts per million (ppm) and D is the dissolved oxygen deficit, equal to the temperature dependent saturation concentration of dissolved oxygen minus the amount of oxygen actually present in the water in ppm. The assumptions inherent in this system of equations and typical values of the parameters are presented in Camp [3], Dobbins [4] and Clark and Viessman [5].

Lee and Hwang [7] have utilized the methods of quasilinearization [6] and invariant imbedding to perform parameter estimation for this formulation. Quasilinearization is an iterative method which considers a linear approximation to the system of ordinary differential equations to the model. Thus some of the models presented in Shastry et al. [1] would have to be linearized. In invariant imbedding, a large number of simultaneous ordinary differential equations is integrated numerically even for a few unknown parameters. The estimator equations are nonlinear even if the original model is linear.

Least squares offers a flexible method which can also be utilized for this problem, and which deals with the differential equations of the model. Like the two other methods, good initial estimates of the parameters are required for convergence on final estimates.

An analytical solution for equation (1) with initial conditions can be obtained for constant R and A which is nonlinear in the parameters k_1, k_2, k_3, R, A. Rather than considering the explicit form of this solution we consider the solution expressed in terms of the eigenvalues and eigenvectors of the system and perturb these to obtain quantities used in the normal equations of regression analysis [8]. Thus, this method is applicable to larger systems of linear ordinary differential equations with constant coefficients. A correction vector to the estimates is calculated iteratively by the modified Newton-Raphson method until the sum of squared errors changes by less than some prescribed amount.

Parametric differentiation of the original system of equations offers an alternate means of obtaining the needed quantities. This method is applicable to nonlinear systems as well as to integro-differential equations [8] and is utilized for identification in cases for which only dissolved oxygen data is available.

As a test case, we consider the determination of k_1, k_2, k_3, R, A and the initial conditions B_o and D_o, from perfect data which is generated by using known values of the parameters. The resultant estimates may then be compared to those used to generate the data both for the perfect data and for deliberately corrupted data. In addition, the method of nonlinear least squares will also be used for parameter estimation for sets of dissolved oxygen data taken under closely controlled laboratory conditions. It is felt that only with data obtained from such conditions can valid parameter identification be performed and the parameters then related to the pertinent physical phenomena.

Least Squares

The minimization of the sum of the squared errors between the observed quantities and a best fit utilizing estimates to the unknown parameters forms the basic idea of least squares. The flexibility and simplicity of least squares makes it one of the most widely used statistical methods. Simple assumptions and modifications of the basic concept allow for maximum likelihood and Bayesian estimates to be obtained.

The mathematical model relating the parameters and the observed quantities is sometimes given by the linear form

$$Y = H\,\theta + U \tag{2}$$

The specific set of data form a sample of the random variables Y and U

$$y = H\,\theta + u \tag{3}$$

for y and N-dimensional vector of data points, θ, a k-dimensional vector for unknown parameters, and H is a N by k matrix of known quantities. Minimization of the sum of squared errors

$$S(\theta) = u^T u \tag{4}$$

for u^T representing the transpose of the vector of errors, u, yields the following estimate to the parameters

$$\theta = [H^T H]^{-1} H^T y \tag{5}$$

Maximum likelihood estimates can be obtained if the errors are assumed to be multivariate normal and if the correlation matrix of the errors, which are assumed to have expected values of zero, is

$$E\ \{UU^T\} = V_u \tag{6}$$

The resulting minimum variance estimates are unbiased

$$\hat{\theta} = [H^T V_u^{-1} H]^{-1}\ H^T V_u^{-1}\ y \tag{7}$$

Although equation (2) indicates that a linear relationship exists between the unknown parameters, it frequently happens that the relationship is a nonlinear one

$$y = h(\theta) + u \tag{8}$$

h can be approximated by a first order Taylor series about initial estimates to the parameters, θ_o

$$y = h(\theta_o) + \sum_{j=1}^{k} \frac{\partial h}{\partial \theta_j}\bigg|_{\theta=\theta_o} \delta\theta_j + u \tag{9}$$

for k unknown parameters. Optimization of $(u^T u)$ with respect to $\delta\theta$ yields the following estimate to the parameters at the ℓ^{th} iteration

$$\hat{\theta}_{\ell+1} = \hat{\theta}_\ell + [H^T H]^{-1}\ H^T\ [y - h(\theta_\ell)] \tag{10}$$

in which

$$H_{ij} = \frac{\partial h_i}{\partial \theta_j} \tag{11}$$

Maximum likelihood estimates are

$$\hat{\theta}_{\ell+1} = \hat{\theta}_\ell + [H^T V_u^{-1} H]^{-1} H^T V_u^{-1}\ [y - h(\theta_\ell)] \tag{12}$$

Bayesian estimates can be obtained if the parameters are assumed to be random variables with a multivariate normal distribution with expected value and covariance matrix of

$$E\{\theta\} = \theta_o \quad E\{(\theta - \theta_o)(\theta - \theta_o)^T\} = V_\theta \tag{13}$$

Maximization of the unconditional likelihood function yields

$$\begin{aligned} \hat{\theta}_{\ell+1} &= \hat{\theta}_\ell + [H^T V_u^{-1} H + V_\theta^{-1}]^{-1} \\ &\{H^T V_u^{-1} [y - h(\hat{\theta}_\ell)] - V_\theta^{-1} [\hat{\theta}_\ell - \theta_o]\} \end{aligned} \tag{14}$$

Often V_u and V_θ are not known a priori and normal least squares yields first estimates to the parameters. A study of the errors can determine whether they are correlated or not. We only consider normal least squares estimates in this work, but should there be additional information, maximum likelihood and

Bayesian estimates can be obtained by the same algorithm, with appropriate modifications. Bayesian formulation for biochemical oxygen demand-dissolved oxygen deficit applies when initial estimates to k_1, k_2, k_3, R, and A have been made from other experiments such as light and dark bottle tests for photosynthesis and winkler tests for BOD.

Eigenvalue Analysis

The mathematical model, equation (1) can be expressed in matrix terms

$$\frac{dx}{dt} = Q\,x + F \tag{15}$$

with initial conditions

$$x_o = \begin{bmatrix} B_o \\ D_o \end{bmatrix} \tag{16}$$

in which

$$X = \begin{bmatrix} B \\ D \end{bmatrix} \quad F = \begin{bmatrix} R \\ -A \end{bmatrix} \quad Q = \begin{bmatrix} -(k_1 + k_3) & 0 \\ k_1 & -k_2 \end{bmatrix} \tag{17}$$

Since this is a linear system the solution is given by

$$x(t) = X(t)x_o + \int_o^t X(t)X^{-1}(\tau)\,F(\tau)\,d\tau \tag{18}$$

where X is the fundamental matrix which satisfies

$$\frac{dX}{dt} = Q\,X \qquad X_o = I \tag{19}$$

with initial conditions equal to the identity matrix.

Since we are considering a linear system with constant coefficients, the solution can be expressed in terms of the eigenvalues and eigenvectors of Q

$$x = Pe^{\Lambda t}P^{-1}x_o + \int_o^t Pe^{\Lambda(t-\tau)}P^{-1}\,F(\tau)\,d\tau \tag{20}$$

for P the matrix of right eigenvectors, P^{-1}, the matrix of left eigenvectors, and Λ the matrix of eigenvalues. Furthermore, for constant F the solution finally reduces to

$$x = Pe^{\Lambda t}P^{-1}\,[x_o + Q^{-1}F] - Q^{-1}F \tag{21}$$

For F, Q, and x_o given by equations (16) and (17), the following are readily calculated

$$e^{\Lambda t} = \begin{bmatrix} e^{-(k_1+k_3)t} & 0 \\ 0 & e^{-k_2 t} \end{bmatrix}$$

$$P = \begin{bmatrix} \dfrac{k_1 + k_3 - k_2}{\sqrt{k_1{}^2 + (k_1 + k_3 - k_2)^2}} & 0 \\ \dfrac{-k_1}{\sqrt{k_1{}^2 + (k_1 + k_3 - k_2)^2}} & 1 \end{bmatrix} \tag{22}$$

$$P^{-1} = \frac{1}{|P|} \begin{bmatrix} 1. & 0 \\ \dfrac{k_1}{\sqrt{k_1{}^2 + (k_1 + k_3 - k_2)^2}} & \dfrac{k_1 + k_3 - k_2}{\sqrt{k_1{}^2 + (k_1 + k_3 - k_2)^2}} \end{bmatrix}$$

where the determinant of P is given by

$$|P| = \frac{k_1 + k_3 - k_2}{\sqrt{k_1{}^2 + (k_1 + k_3 - k_2)^2}} \tag{23}$$

and the solution becomes

$$B = [B_o - \frac{R}{k_1 + k_3}]\ e^{-(k_1+k_3)t} + \frac{R}{k_1 + k_3}$$

$$D = e^{-k_2 t}\ \{D_o + \frac{[\frac{Rk_1}{k_1 + k_3} - A]\ [e^{k_2 t} - 1]}{k_2} \tag{24}$$

$$+ k_1 \frac{[B_o \frac{-R}{k_1 + k_3}]\ [e^{(k_2-k_1-k_3)t} - 1]}{k_2 - k_1 - k_3} \}$$

Thus for observed data in the form of biochemical oxygen demand and dissolved oxygen deficit, and unknown parameters, k_1, k_2, k_3, R, A, B_o, and D_o; we have at the i^{th} data point

$$B_i = B(t_i,\theta) + u_{b_i}$$
$$D_i = D(t_i,\theta) + u_{d_i} \tag{25}$$

where B and D are clearly nonlinear in the seven unknowns.

The partial derivatives of the BOD and DO with respect to the unknown parameters are required for the normal equations, equation (10), of regression analysis. These could be obtained by differentiating the solution given by equation (24). Rather than doing this, however, we will present a method based on the perturbation of the eigenvectors and eigenvalues of the system. This is general and can be applied to larger systems of linear ordinary differential equations with constant coefficients.

The partial derivative of the solution given by equation (21) with respect to the parameters is expressed in terms of the partial derivative of the system matrix, Q, the matrix of right eigenvector, P and left eigenvectors, P^{-1}, the eigenvalue matrix, Λ, and the forcing vector, F, with respect to the parameter, θ

$$\begin{aligned}\frac{\partial x}{\partial \theta} &= \frac{\partial}{\partial \theta}[Pe^{\Lambda t}P^{-1}][x_o + Q^{-1}F] + Pe^{\Lambda t}P^{-1}\frac{\partial x_o}{\partial \theta} \\ &+ [Pe^{\Lambda t}P^{-1} - I]\frac{\partial}{\partial \theta}[Q^{-1}F]\end{aligned} \tag{26}$$

The partial derivative of Q^{-1} and P^{-1} with respect to the parameters can be expressed in terms of the partial derivative of Q and P respectively [8]

$$\begin{aligned}\frac{\partial Q^{-1}}{\partial \theta} &= -Q^{-1}\frac{\partial Q}{\partial \theta}Q^{-1} \\ \frac{\partial P^{-1}}{\partial \theta} &= -P^{-1}\frac{\partial P}{\partial \theta}P^{-1}\end{aligned} \tag{27}$$

since

$$Q^{-1}Q = P^{-1}P = I \tag{28}$$

For right eigenvectors, u_k, the optimization of the quadratic form

$$v_k^T\,[Q-\lambda_k I]\,u_k \tag{29}$$

for v_k the left eigenvector, yields the perturbation of the eigenvalue with respect to θ

$$\frac{\partial \lambda_k}{\partial \theta} = v_k^T\frac{\partial Q}{\partial \theta}u_k \tag{30}$$

For u_k the normalized eigenvector associated with λ_k (u_k is the k^{th} column of the matrix P)

$$u_k^T u_k = 1 \tag{31}$$

the perturbation of the eigenvector can be expressed as a summation of n − 1 terms, (n is the dimension of Q, two for our problem)

$$\frac{\partial u_k}{\partial \theta} = \sum_{i \neq k}^{n} \gamma_{ik} v_i \tag{32}$$

Optimization of the quadratic forms

$$v_\ell^T (Q - \lambda_k I) u_k \tag{33}$$

yields n − 1 equations for the n − 1 unknown constants, γ_{ik}

$$v_\ell^T \frac{\partial Q}{\partial \theta} u_k + (\lambda_\ell - \lambda_k) \sum_{i \neq k}^{n} \gamma_{ik} v_\ell^T v_i = 0 \quad \ell \neq k \tag{34}$$

The partial derivatives of Q, $e^{\Lambda t}$, and P are readily obtained from equations (17), (30), (32), and (34). For k_1 the unknown parameter, for example, there results,

$$\begin{aligned} \frac{\partial Q}{\partial k_1} &= \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} \\ \frac{\partial e^{\Lambda t}}{\partial k_1} &= \begin{bmatrix} -te^{-k_1 t} & 0 \\ 0 & 0 \end{bmatrix} \\ \frac{\partial P}{\partial k_1} &= \frac{k_2 - k_3}{[k_1{}^2 + (k_1 + k_3 - k_2)^2]^{3/2}} \begin{bmatrix} k_1 & 0 \\ k_1 + k_3 - k_2 & 0 \end{bmatrix} \end{aligned} \tag{35}$$

which are identical to values obtained by differentiation equation (22). Since Q is not symmetric we are not assured that the eigenvalues and associated eigenvectors are real. Our analysis will be restricted to real eigenvalues, and thus exponential behavior, however. For complex eigenvalues the solution to the homogeneous system of equations is made up of decaying sinusoids [8].

This method is restricted to cases where the differential equations are linear with constant coefficients. However, it is applicable to systems larger than two. Nevertheless, a method is required which can deal with nonlinear models such as those given in Shastry, et al. [1].

Parametric Differentiation

Another method of obtaining the partial derivatives with respect to the parameters considers the system of differential equations rather than the solution to the system. Thus it is applicable to nonlinear systems as well as to linear ones. The sensitivity equations are obtained by taking the partial derivative of the system with respect to the parameter and interchanging differentiation in time and the parameter. Thus for equation (15),

$$\frac{d}{dt}\left(\frac{\partial x}{\partial \theta}\right) = Q \frac{\partial x}{\partial \theta} + \frac{\partial Q}{\partial \theta} x + \frac{\partial F}{\partial \theta} \tag{36}$$

These can be solved either by numerical integration or by eigenvalue analysis if the F vector is of simple form, such as sinusoidal or constant functions.

In some situations of water quality observation, the dissolved oxygen is observed but the biochemical oxygen demand is not. A method of parameter estimation based on this data is required. For non-constant R in equation (15), the analytical solution for the biochemical oxygen demand can be obtained by the integrating factor technique

$$B = B_o\, e^{-(k_1+k_3)t} + \int_o^t e^{-(k_1+k_3)(t-\tau)}\, R(\tau)\, d\tau \tag{37}$$

and the equation for the dissolved oxygen deficit becomes

$$\frac{dD}{dt} + k_2 D = -A(t) + k_1 \quad \{B_o e^{-(k_1+k_3)t} + \int e^{-(k_1+k_3)(t-\tau)}\, R(\tau)d\tau\} \tag{38}$$

Parametric differentiation of this equation with respect to the unknown θ yields the desired sensitivity equation

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial D}{\partial \theta}\right) + k_2 \frac{\partial D}{\partial \theta} &= \frac{\partial A(t)}{\partial \theta} - D \frac{\partial k_2}{\partial \theta} + \frac{\partial}{\partial \theta}\left[k_1\, B_o e^{-(k_1+k_3)t}\right] \\ &+ \frac{\partial k_1}{\partial \theta} \int_o^t e^{-(k_1+k_3)(t-\tau)}\, R(\tau)d\tau \\ &+ k_1 \int_o^t \frac{\partial}{\partial \theta}\left[e^{-(k_1+k_3)(t-\tau)}\, R(\tau)\right] d\tau \end{aligned} \tag{39}$$

The partial derivatives of D with respect to k_1, k_2, k_3, B_o, and D_o can be obtained, as well as with respect to R and A should these be constant.

Numerical Results

The methods of eigenvalue analysis and parametric differentiation are utilized on simulated data with and without generated noise with zero mean and standard deviation .1. These are utilized to obtain estimates for which we know the answer, and will indicate the applicability of the methods to perfect data and data with errors which are independent and Gaussian. For cases one and two, the method of eigenvalue perturbation is used for estimation of the seven parameters for the situation in which BOD and D are observed, and for which only dissolved oxygen data is available. Eigenvalue perturbation and parametric differentiation are used to obtain estimates for k_1, k_2, A and D_o from observation of oxygen data only in cases three and four. The fifth case considers determination of all seven parameters, from oxygen data only, by parametric differentiation.

Iteration is performed until the sum of the squared errors changes by less

Table 1.

Parameter	$k_1 (day^{-1})$	$k_2 (day^{-1})$		$R (ppm/day)$	$A (ppm/day)$	$B_o (ppm)$	$D_o (ppm)$	$Var (ppm^2)$	Iterations
Case[a]									
Value	.75	1.50	.025	.15	.85	5.0	1.0	—	—
Initial Est.	.70	1.45	.030	.15	.80	4.8	.9	—	—
1	.75	1.50	.025	.15	.85	5.0	1.0	.79E-12	4
	(.74E-12)	(.45E-11)	(.84E-12)	(.18E-12)	(.27E-12)	(.11E-12)	(.21E-12)		
1σ[b]	.74	1.50	.002	.08	.83	4.95	1.00	.87E-2	4
	(.85E-2)	(.53E-1)	(.95E-2)	(.19E-2)	(.29E-2)	(.11E-2)	(.23E-2)		
2	.70	1.50	.07	–.14	.58	4.94	1.00	.77E-12	5
	(–.96E-9)	(.21E-9)	(–.82E-9)	(–.83E-8)	(–.62E-8)	(–.20E-7)	(.25E-12)		
2σ[b]	.57	3.11	–.333	–.2.66	–1.36	5.63	.94	1.01E-2	25
	(–114.)	(1.89)	(–114.)	(14024.)	(–30337.)	(–42573.)	(.41E-2)		
3[c]	.75	1.50	—	—	.85	—	1.00	.18E-11	4
	(.50E-12)	(.39E-11)	—	—	(.15E-11)	—	(.44E-12)		
3σ[b,c]	.73	1.46	—	—	.88	—	1.00	76E-2	15
	(.22E-2)	(.18E-1)	—	—	(.64E-2)	—	(.19E-2)		
4[c]	.75	1.50	—	—	.85	—	1.00	.90E-12	4
	(.26E-12)	(.20E-11)	—	—	(.76E-12)	—	(.22E-12)		
4σ[b,c]	.73	1.46	—	—	.88	—	1.00	.76E-2	6
	(.22E-2)	(.18E-1)	—	—	(.64E-2)	—	(.19E-2)		
5	.75	1.50	.022	–.01	.69	4.77	1.00	.80E-10	7
	(–.13E-6)	(.23E-7)	(.11E-6)	(–.11E-5)	(–.11E-5)	(–.17E-5)	(.27E-10)		
5σ[b]	1.12	3.48	–.91	–1.73	–1.74	3.14	.93	81E-2	25
	(–32.0)	(1.33)	(–32.1)	(1707.3)	(27420.)	(18884.)	(.35E-2)		

[a] The ten cases consisted of data with 101 data points over 3 days.

[b] σ Corresponds to corrupted data with error having mean of zero and standard deviation equal to .10.

[c] In case 3 and 4 the parameters k_3, R, and B_o were set at correct values.

than .00001 or until twenty-five iterations are required. For cases in which both BOD and D are observed, each data point is weighed equally. Initial estimates are given in the table. For cases in which k_3, R, and B_0 are not determined, these values are set equal to the values used to generate the data. Subroutines GAUSS and RANDU of the IBM Scientific Subroutine Package are used to generate the errors on an IBM 360/95.

The results of these five cases are given in Table 1. The correct values, initial estimates, and final estimates for the parameters are shown. The variance of the errors at the final estimate is indicated for each case as is the number of iterations required for convergence. The variance of the parameters are shown in the table underneath the parameter in parentheses. The best fit lines and the generated data for BOD(+) and dissolved oxygen deficit(X) are given in Figure 1 where BOD has been nondimensionalized by the initial BOD equal to 5 ppm and the dissolved oxygen data (saturated minus deficit) is nondimensionalized by the saturated value set equal to 5 ppm for convenience. Lines denoted by 2 and 2σ correspond to results from uncorrupted and corrupted data respectively, for case 2. The corrupted data and best fits are shown in Figure 2.

The fits to the dissolved oxygen data are good for both the corrupted and perfect data for all five cases. The best fit to the dissolved oxygen data is good

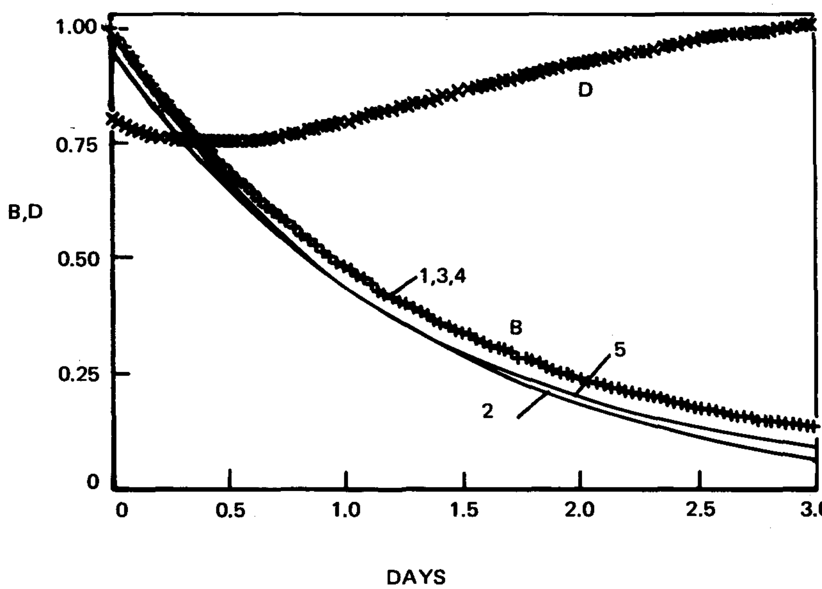

Figure 1.

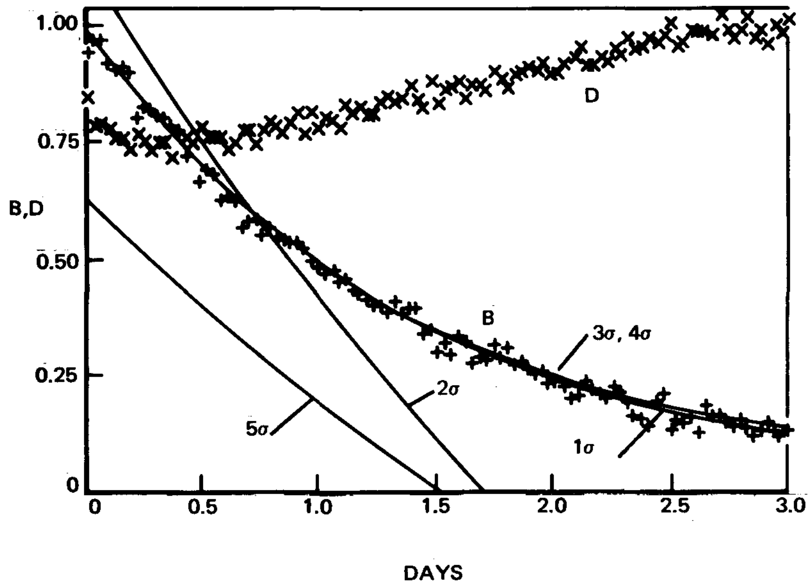

Figure 2.

for case 1 and 1σ which considered this data but is not very good for all other cases. For cases 3 and 4 which assumed k_3, R and B_0 as known, it is better than cases 2 and 5 which treated these as unknowns.

The estimates of the parameters for uncorrupted data are the same for eigenvalue perturbation and parametric differentiation methods in cases 3 and 4 and 3σ and 4σ, but differ between case 2 and 5 and 2σ and 5σ. When an estimate is not very good there is a large variance associated with it indicating the merit of the estimates. From these results it is shown that good estimates may be obtained if both BOD and D data are available, or for D data only, if k_3, R, and B_0 are known.

Experimental Study

A laboratory scale experiment was conducted in which natural reaeration phenomena was examined through a wide range of dynamic conditions. In particular the experimental study was focused on the influence of wind action upon reaeration phenomena.

The facilities used for this experimental investigation were those of the Institut de Mecanique Statistique de la Turbulence, directed by Professor A. Favre in Marseilles, France. The wind tunnel-water channel facility had a test section consisting of an 8m fetch wherein wind velocities could be varied through the range from 0-15 m/sec. and where water velocities varied from 0-18 cm/sec in a channel 15 cm deep by 55 cm wide. In both the air and water flows, constant temperature (20°C) conditions were maintained via heat exchanges in both closed fluid circuits. Further details concerning the experimental facility can be found in Reference 9.

The experimental procedure consisted of chemically extracting the dissolved oxygen from the water. This extraction used catalyzed sodium sulfide and reduced the dissolved oxygen deficit to 10 or 20% of the saturation value. Each individual experiment then consisted of establishing the desired temperature and velocities for both the air and the water and then monitoring the dissolved oxygen increase to saturation conditions using Winkler and osmosis methods. In all of these laboratory scale experiments, no BOD data were recorded.

The dissolved oxygen results for constant water velocity with variable wind velocity conditions are presented via the data points in Figure 3. Fitting these dimensionless dissolved oxygen uptake curves using the nonlinear least squares techniques with parametric differentiation described above gives the constants tabulated in Table 2 for constant water velocity. The best fit lines are shown graphically in Figure 3 and in tabular form in Table 2.

Comparing the reaeration coefficient, k_2 as presented in Table 2 for the case

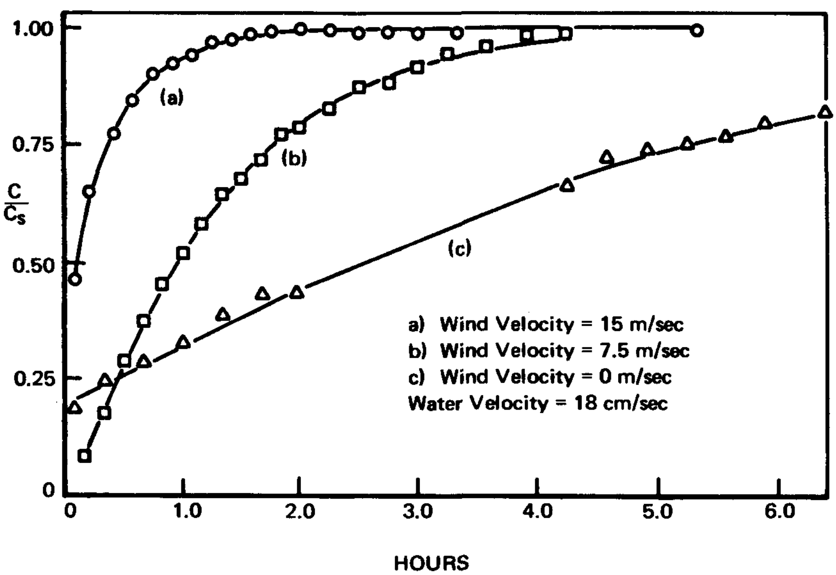

Figure 3.

Table 2. Parameters for Best Fit Lines (Values in parentheses are variances)

	Water velocity (cm/sec)	*Air velocity (m/sec)*	k_1 (Hrs^{-1})	k_2 Hrs^{-1})	B_o (ppm)	D_o (ppm)	*Var*
case a	18.0	15.0	1.9971 (0.0259)	8.0504 (0.0259)	10.45	5.5050 (0.0333)	0.00366
case b	18.0	7.5	0.9163 (0.0008)	1.9374 (0.0002)	10.80	7.5189 (0.01444)	0.00532
case c	18.0	0.0	0.2816 (0.0008)	0.6229 (0.0004)	10.00	5.7601 (0.0306)	0.0390

of no wind to that for wind velocity of 15 m/sec it is found that a factor of approximately 13 is found. It is concluded that the wind can exert a significant influence on natural reaeration.

Conclusions

The method of least squares together with eigenvalue perturbation and parametric differentiation are shown to yield good estimates of water quality parameters for observation of both the biochemical oxygen demand and the dissolved oxygen deficit. The method of parametric differentiation can also be used for parameter estimation for dissolved oxygen data only if the initial BOD, the sedimentation rate, and the rate of runoff are known.

These methods are utilized on corrupted and uncorrupted data obtained by simulation of the time series. Parametric differentiation is utilized on actual physical data of dissolved oxygen only.

Although these methods are applied to a simple model, they may, in general, be used for more complicated models should the data require this. Second order terms for the decay rates could be considered and the effect of nitrification could be modeled, thus increasing the number of equations. Parametric differentiation could also be used for nonlinear models such as those considered by Shastry, Fan and Erickson [1]. Although R and A are considered constant here, they could vary along the stream or be random variables. Again this method could be used along with numerical integration of both the equations for D and BOD and the sensitivity equation. Finally, partial differential equations could be treated using this method of parametric differentiation.

REFERENCES

1. J. S. Shastry, L. T. Fan, and L. E. Erickson, Nonlinear parameter estimation in water quality modeling, *Jnl. of the Environmental Engr.*, Div. ASCE, June 1973.

2. R. W. Cleary and D. D. Adrian, New analytical solutions for dye diffusion equations, *Jnl. Environmental Engr.*, Div. ASCE, June 1973.
3. T. R. Camp, *Water and Its Impurities*, Reinhold Publishing Corporation, New York, 1963.
4. W. E. Dobbins, BOD and oxygen relationships in streams, *Jnl. Sanitary Engr.*, Division ASCE, 90, 1964.
5. J. W. Clark and W. Veissman, Jr., *Water Supply and Pollution Control*, International Textbook Company, Scranton, Pa., 1965.
6. E. S. Lee and I. Hwang, Stream quality modeling by quasilinearization, *Jnl. Water Pollution Control Federation*, Feb. 1971.
7. E. S. Lee and I. Hwang, Dynamic modeling of stream quality by invariant imbedding, *Water Resources Bulletin*, Feb. 1971.
8. J.-G. Beliveau, *Suspension Bridge Aeroelasticity: Nonlinear Least Squares Techniques for System Identification*, Ph.D. Thesis, Dept. of Civil and Geological Engineering, Princeton University, Sept. 1973.
9. M. Coantic and A. Favre, Air Sea Interactions: Research Program and Facilities at I.M.S.T., Eighth Symposium on Naval Hydrodynamics, ACR 179, ONR 1970.