Biodegradation (2010) 21:203–216
DOI 10.1007/s10532-009-9294-7

ORIGINAL PAPER

Novel differential elimination method for determining kinetic coefficients under substrate self-inhibition

Seongjun Park · Bruce E. Rittmann · Wookeun Bae

Received: 22 October 2008 / Accepted: 18 August 2009 / Published online: 19 September 2009

Abstract A differential elimination method (DEM) is developed to determine the kinetic coefficients for substrate self-inhibition. Finite differentiation of the equation eliminates either K_I or K_S, which enables the equation to be linearized so that $\hat{q}$, K_S, and K_I can be estimated without using nonlinear least square regression (NLSR). The DEM options that eliminate K_I or K_S computed the parameter values exactly when the data did not contain any errors. If one-point or random errors were not too large, both DEM options worked as well as NLSR when data were acquired with geometric intervals for substrate concentration. The DEM was more accurate for fitting the data for the smallest and largest values of S, but relatively weaker in estimating the observed maximum substrate utilization rate, q_{max}. The estimates for S_{max}, the concentration at which the maximum specific substrate utilization rate is observed, were relatively invariant among the methods, even when K_S and K_I differed. When the intervals were arithmetic (i.e., equal intervals of substrate concentration) and the data contained errors, the DEM and NLSR estimated the parameters poorly, indicating that collecting data with an arithmetic interval greatly increases the risk of poor parameter estimation. Parameter estimates by DEM fit very well experimental data from nitrification or photosynthesis, which were taken with geometric intervals of substrate concentration or light intensity, but fit poorly phenol-degradation data, which were obtained with arithmetic substrate intervals. Besides providing a reasonable substitute for NLSR, the DEM also can be used as a tool to diagnose the quality of experimental data by comparing its estimates between the DEM options, or, more rigorously, to those from NLSR.

Keywords Differential elimination method · Graphical plot method · Kinetic coefficients · Linear plot method · Substrate inhibition · Nonlinear least square regression

S. Park · B. E. Rittmann
Center for Environmental Biotechnology, Biodesign Institute at Arizona State University, 1001 South McAllister Avenue, Tempe, AZ 85287-5701, USA
e-mail: Seongjun.Park@asu.edu

B. E. Rittmann
e-mail: Rittmann@asu.edu

W. Bae (✉)
Department of Civil and Environmental Engineering, Hanyang University, Sa 1-Dong, Ansan, Gyunggi-Do, Republic of Korea
e-mail: wkbae@hanyang.ac.kr

Introduction

Substrate self-inhibition, the slowing of the metabolic reaction when the substrate concentration is high, has been observed frequently in the treatment of wastewaters with high concentrations of phenol or

ammonium. It is also observed for light intensity with cell growth during photosynthesis (Wang et al. 1999). Substrate self-inhibition has been studied by many researchers (e.g., Andrews 1968; Haldane 1930; Park and Bae 2009; Vadivelu et al. 2006; Wang et al. 1999), and Eq. 1 shows the generally agreed upon kinetic expression.

$$q = \frac{\hat{q}S}{K_S + S + \frac{S^2}{K_I}} = \frac{\frac{\hat{q}}{(1+S/K_I)}S}{\frac{K_S}{(1+S/K_I)} + S} \quad (1)$$

in which, q and $\hat{q}$ are a specific substrate utilization rate and its maximum value (mgS/mgVSS-d; VSS = volatile suspended solids), respectively, S is a substrate concentration (mgS/L), K_S is a half-maximum-rate concentration (mgS/L), and K_I is the self-inhibition concentration (mgS/L).

The right-hand expression in Eq. 1 emphasizes that the maximum specific substrate utilization rate (numerator) and the half-maximum-rate concentration (denominator) are affected by the substrate concentration, a variable. Therefore, simple linearization methods to find $\hat{q}$and K_S for the Monod model, such as the Lineweaver–Burk treatment, cannot be used to determine the kinetic coefficients with self-inhibition. Instead, complicated nonlinear regression methods have been adopted (e.g., Park and Bae 2009; Chen et al. 2005; Futamata et al. 2001; Goudar et al. 2000).

The objective of this research is to develop a simple, but mathematically sound linearization method to estimate $\hat{q}$, K_S, and K_I for substrate self-inhibition. We test the method with artificially generated data sets to see how it responds to various types of errors in the data. We also test it using experimental data obtained from the literature for nitrification, photosynthesis, and phenol degradation. In each case, we compare the results of our simple method with estimates by nonlinear least-square regression.

Method development

Existing graphical linear-plot methods

A commonly used approach for estimating kinetic parameters for enzymes or whole cells is a linearization of the hyperbolic equation. A classic example is the Lineweaver–Burk (LB) treatment (Lineweaver and Burk 1934) of taking reciprocals of the Michalis-Menton or Monod substrate-utilization equation.

$$\frac{1}{q} = \frac{K_S 1}{\hat{q}\ S} + \frac{1}{\hat{q}} \quad (2)$$

where q and S are input from experimental data. The values of $\hat{q}$ and K_S are estimated from the *y*-intercept and the slope, respectively. Estimates by the LB method are affected most seriously from errors for low q values, because the errors are amplified by using the reciprocal.

The Hanes-Woolf (HW) method is a modification of the LB method in which all terms in the LB equation are multiplied by the substrate concentration, S (Bailey and Ollis 1986).

$$\frac{S}{q} = \frac{K_S}{\hat{q}} + \frac{S}{\hat{q}} \quad (3)$$

In comparison with the LB method, the HW method has an advantage of reducing the amplification effect of experimental errors for low q values, but a disadvantage of amplifying errors for high substrate concentrations.

Linearization by taking reciprocals cannot be applied directly with self-inhibition. An LB-type reciprocal treatment of Eq. 1 (Eq. 4) or its HW modification (Eq. 5) is not a linear function:

$$\frac{1}{q} = \frac{K_S 1}{\hat{q}\ S} + \frac{1}{\hat{q}} + \frac{S}{\hat{q}K_I} \quad (4)$$

$$\frac{S}{q} = \frac{K_S}{\hat{q}} + \frac{S}{\hat{q}} + \frac{S^2}{\hat{q}K_I} \quad (5)$$

An approximation method that relaxes the nonlinearity problem is possible when S is very small, because the $S/\hat{q}K_I$ term in Eq. 4 (or $S^2/\hat{q}K_I$ in Eq. 5) becomes insignificant. Then, the equations simplify to the original reciprocal equations for no inhibition. The serious shortcomings with the approximation are that it often is very difficult to obtain reliable experimental data when S is very small, which makes the method impractical, and that K_I cannot be estimated reliably this way, as the effects of inhibition are negligible for very small S.

Another linearization technique involves substituting K_I with S^2_{max}/K_S, based on the fact that the "optimal" substrate concentration, which gives the maximum reaction rate, is expressed as $S_{max} = \sqrt{K_S K_I}$

(Wang et al. 1999; Kotturi et al. 1991). Thus, a linear equation is obtained from Eq. 4.

$$\frac{1}{q} = \frac{1}{\hat{q}} + \frac{K_S}{\hat{q}}\left(\frac{1}{S} + \frac{S}{S_{max}^2}\right) \tag{6}$$

where $(1/S + S/S_{max}^2)$ is the independent variable. The S_{max} value must be estimated independently from a set of experimental data for q versus S (Wang et al. 1999), and the accuracy of this method depends on the accuracy in the estimation of the S_{max}.

Linearization by differential elimination method (DEM)

We propose a simple and mathematically sound differential elimination method (DEM) to obtain kinetic parameters for substrate self-inhibition. The method linearizes the self-inhibition equation by eliminating either K_I (called the K_I-elimination option) or K_S (the K_S-elimination option). The remaining two parameters ($\hat{q}$ and K_S in K_I-elimination option, or $\hat{q}$ and K_I in K_S-elimination option) are graphically estimated using experimental data. Then, the third parameter can be determined with Eqs. 4 or 5.

For K_I-elimination, the LB-type reciprocal expression in Eq. 4 can be divided by S to give

$$\frac{1}{qS} = \frac{K_S}{\hat{q}S^2} + \frac{1}{\hat{q}S} + \frac{1}{\hat{q}K_I} \tag{7}$$

Taking finite differences for Eq. 7 gives

$$\begin{aligned}\frac{1}{q_iS_i} - \frac{1}{q_{i+1}S_{i+1}} &= \left(\frac{K_S}{\hat{q}S_i^2} + \frac{1}{\hat{q}S_i} + \frac{1}{\hat{q}K_I}\right) \\ &\quad - \left(\frac{K_S}{\hat{q}S_{i+1}^2} + \frac{1}{\hat{q}S_{i+1}} + \frac{1}{\hat{q}K_I}\right) \\ &= \frac{K_S}{\hat{q}}\left(\frac{1}{S_i^2} - \frac{1}{S_{i+1}^2}\right) + \frac{1}{\hat{q}}\left(\frac{1}{S_i} - \frac{1}{S_{i+1}}\right) \\ &= \left(\frac{K_S}{\hat{q}}\left(\frac{S_{i+1} + S_i}{S_iS_{i+1}}\right) + \frac{1}{\hat{q}}\right)\left(\frac{S_{i+1} - S_i}{S_iS_{i+1}}\right)\end{aligned} \tag{8}$$

A rearrangement of Eq. 8 by dividing both sides by $(S_{i+1} - S_i)/(S_iS_{i+1})$ gives

$$\frac{S_i/q_{i+1} - S_{i+1}/q_i}{(S_i - S_{i+1})} = \frac{K_S}{\hat{q}}\left(\frac{1}{S_i} + \frac{1}{S_{i+1}}\right) + \frac{1}{\hat{q}} \tag{9}$$

Equation 9 has eliminated K_I and is linearized with the independent parameter $(1/S_i + 1/S_{i+1})$. The $\hat{q}$ and K_S values can be found from the intercept and slope of Eq. 9, respectively. Equation 8 resembles Eq. 2, but is more complicated because the dependant variable involves finite differences. Since the q values are taken as reciprocals, an amplification of the error in low q values may take place, as it does in the traditional LB treatment.To determine K_I, Eq. 4 can be rearranged to

$$\frac{\hat{q}}{q} - \frac{K_S}{S} - 1 = \frac{S}{K_I} \tag{10}$$

In Eq. 10, S in the right-hand term is an independent variable, whereas the left-hand terms constitute a dependent variable. With $\hat{q}$ and K_S known from Eq. 9, K_I can be estimated from the slope of a graphical presentation of the data with Eq. 10.

A similar treatment can be applied to the HW-type reciprocal equation (Eq. 5) for K_S-elimination.

$$\begin{aligned}\frac{S_i}{q_i} - \frac{S_{i+1}}{q_{i+1}} &= \left(\frac{K_S}{\hat{q}} + \frac{S_i}{\hat{q}} + \frac{S_i^2}{\hat{q}K_I}\right) - \left(\frac{K_S}{\hat{q}} + \frac{S_{i+1}}{\hat{q}} + \frac{S_{i+1}^2}{\hat{q}K_I}\right) \\ &= \left(\frac{1}{\hat{q}K_I}(S_i + S_{i+1}) + \frac{1}{\hat{q}}\right)(S_i - S_{i+1})\end{aligned} \tag{11}$$

Rearranging leads to

$$\frac{S_i/q_i - S_{i+1}/q_{i+1}}{(S_i - S_{i+1})} = \frac{1}{\hat{q}K_I}(S_i + S_{i+1}) + \frac{1}{\hat{q}} \tag{12}$$

This linearization gives $\hat{q}$ from the *y*-intercept and K_I from the slope when the independent variable is $(S_i + S_{i+1})$. Equation 12 is similar to Eq. 9 in that the q values are in a reciprocal form. Again, the third unknown, K_S, can be determined by rearranging Eq. 4 to

$$\frac{\hat{q}}{q} - \frac{S}{K_I} - 1 = K_S\frac{1}{S} \tag{13}$$

We evaluate the K_I-eliminated approach (Eqs. 9, 10) and the K_S-eliminated approach (Eqs. 12, 13) with artificially generated data and experimental data. We compare the results from each approach with each other and against those from nonlinear least square regression.

Materials and methods

Error test for differential elimination method

We investigated the ability of the differential elimination method (DEM) to extract parameter values from artificially generated data sets having characteristic types of errors. First, two sets of error-free data for q versus S were generated by setting 5 mgS/mgVSS-d for $\hat{q}$, 10 mgS/L for K_S, and 100 mgS/L for K_I. We used six pairs of data points, which is typical for experimental data. The intervals of the independent variable (substrate concentration) were assigned in two ways: geometric or arithmetic. The values of q for the test values of S are given in Table 1.

We then introduced errors to the data sets using three characteristic types: uniform, one-point, or random. Uniform errors were generated by multiplying the calculated true q values (Table 1) by 1.1, corresponding to +10% errors. One-point error was applied to a single datum, which was a q value that corresponded to the lowest, a middle (the third from the beginning), or the highest value of S. Random errors were generated randomly, but maintaining the normal distribution property of the errors. Four errors in the range of 0 to ±5% of the true value (0, −1, 3, and 5%) and two in the range of ±5 to ±10% (−5 and 7%) were generated randomly using the 'Matlab ver. 7.0' program. The results roughly correspond to a normal distribution of errors whose standard deviation was 5% ($\sigma = \pm 5\%$). Then, the errors were imposed upon the q values one each in a random manner. For a thorough test, 15 sets of data with random errors, but in different sequences, were generated. Two of the data sets are shown in Table 2 as examples. Additional extensive data sets with errors (not shown) also were generated and tested as needed to see the performance and characteristics of the DEM. Particularly, we generated 3,000 data sets (1,000 data sets for each error level, 5, 10, and 15%) for testing DEM with six points having geometric intervals, utilizing the function of Matlab 7.0, 'randn', which generates random data from a standard normal distribution curve. Then, we prepared the normally distributed error data with the conditions of mean (on average for all sets) = 0 and $\sigma = 5$, 10, and 15%. Although we did not adjust each set of data to have mean = 0, the mean of errors of the whole 1,000 data sets approached 0. For example, the mean of errors for the 1,000 data sets with 15% error level was 0.09%. The 15% error level provided several extreme cases, in which a point error was as high as 54%. The data with extreme errors tested the performance of DEM and nonlinear least square regression (NLSR) rigorously. These added data sets are further explained as the results are presented.

Table 1 Error-free data sets for substrate inhibition with two different schemes of data-acquisition intervals

Geometric interval of S		Arithmetic interval of S	
Concentration, S (mgS/L)	Specific rate, q (mgS/mgVSS-d)	Concentration, S (mgS/L)	Specific rate, q (mgS/mgVSS-d)
1	0.45	40	3.03
3	1.15	80	2.60
9	2.27	120	2.19
27	3.05	160	1.88
81	2.59	200	1.64
243	1.44	240	1.45

$\hat{q}$: 5 mgS/mgVSS-d, K_S:10 mgS/L, and K_I: 100 mgS/L

Table 2 Two examples of data sets with random errors on q in geometric intervals

Error set I			Error set II		
Concentration, S (mgS/L)	Specific rate, q (mgS/mgVSS-d)	Error (%)	Concentration, S (mgS/L)	Specific rate, q (mgS/mgVSS-d)	Error (%)
1	0.45	−1	1	0.43	−5
3	1.09	−5	3	1.20	5
9	2.34	3	9	2.43	7
27	3.20	5	27	3.14	3
81	2.77	7	81	2.59	0
243	1.44	0	243	1.43	−1

Selection of experimental data from the literature

We tested the performance of the DEM with three sets of experimental data from the literature. The experiments were on nitrification, photosynthesis, and phenol biodegradation, which are known to exhibit substrate self-inhibition. Among the data sets, the nitrification and photosynthesis experiments adopted a geometric interval for substrate concentration, whereas the phenol biodegradation experiment used an arithmetic interval. We compared the parameter values found from the DEM to those from nonlinear least square regression (NLSR).

The nitrification data, obtained from Park and Bae (2009), involved batch oxidation of ammonium (initial concentrations: 5–450 mgN/L) by a pre-acclimated mixed-culture sludge at constant temperature and pH. In the case of photosynthesis, Wang et al. (1999) cultivated a cyanobacterial strain, *Synechococcus* sp., with medium containing ethanol. Various light intensities were applied by tungsten lamps ranging from 1 to 20 klux. The strain was genetically modified such that ethylene was produced during photosynthesis. The ethylene concentration in the exhaust gas from the vessel was measured by a gas chromatography. The specific production rate of ethylene as a function of light intensity showed a typical substrate self-inhibition pattern. For phenol biodegradation, data were obtained from Goudar et al. (2000), who used mixed-culture sludge in batch experiments with initial phenol concentrations from 100 to 1,300 mg/L. They concluded that the Andrews self-inhibition model was the best suitable among nine different kinetic models.

The data from Park and Bae (2009) were taken from the original data files. The data of photosynthesis and phenol biodegradation were estimated from the figures in the papers of Wang et al. (1999) and Goudar et al. (2000), respectively. The graphs were scanned by using HP C4128 with 600dpi resolution, each image was imported to Origin pro 7.5 (Microcal software, Inc.), and the coordinates of each datum were read by a PC from the scanned image.

Nonlinear least-square regression (NLSR)

We adopted NLSR as a reference to evaluate the results from the DEM. The concept of NLSR is based on the partial differentiation equation of each kinetic parameter and a matrix technique such as Gaussian elimination (Bates and Watts 1988). Equation 1 is adapted into the general statistical model ($y = f(x_i, \beta) + \varepsilon$) assumed for the NLSR problems. y denotes the response (dependent) variable, x denotes predictor (independent) variables, β (unsubscripted) denotes the vector of unknown parameters to be estimated, and ε is an estimator value. We used the function for nonlinear regression in Origin pro 7.5 (Microcal software, Inc.) for the calculation by NLSR. After Eq. 1 was created in the function database, data were input to the spreadsheet, and the nonlinear regression function was run to best fit the three dependent variables.

Results and discussion

Responses to artificial errors

With error-free or uniform-error data

The DEM, as well as NLSR, estimated perfectly accurate parameters with the error-free data given in Table 1. K_I-elimination, K_S-elimination, and NLSR gave exactly the same (and correct) parameter estimates using the data with geometric interval and arithmetic intervals.

When 10% uniform errors were imposed upon the q values in the data, the DEM and NLSR estimated exactly the same parameters again. This time, the estimates for K_S and K_I were correct, but the estimated $\hat{q}$ was 10% higher than the true value, since each q value contained +10% errors.

With one-point error

The linearization results with K_S-elimination for data containing one point error are shown in Fig. 1 (for geometric-interval data) and Fig. 2 (for arithmetic-interval data). The differential elimination equations gave, in general, higher R^2 values and less systematic errors with the geometric-interval data than with the arithmetic-interval data. The K_I-elimination option gave similar results (data not shown).

All results are compiled in Table 3, including those from NLSR. In general, one point error with arithmetic-interval data made it impossible to estimate the parameters accurately with any method. The estimated

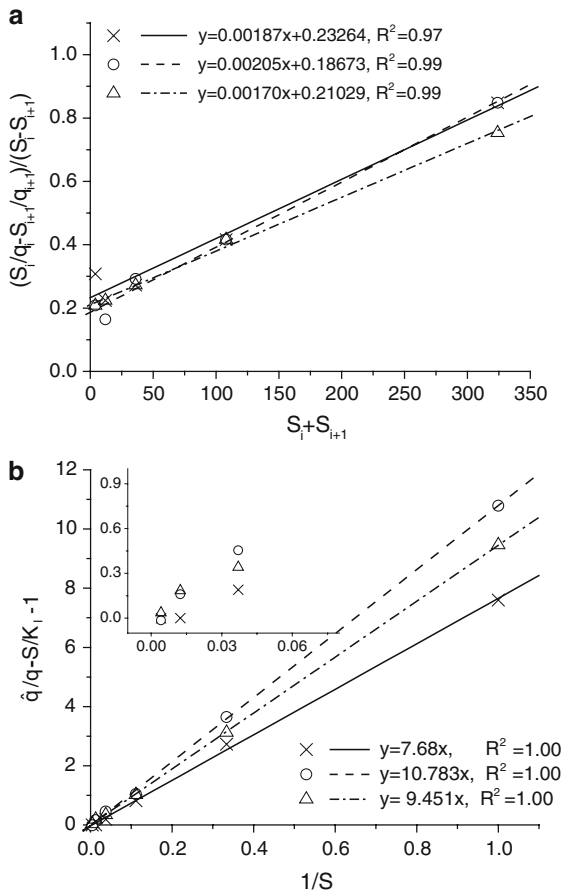

Fig. 1 Influence of one point error with geometric-interval data using the K_S-elimination option. (×: point error at 1 mgS/L, ○: point error at 9 mgS/L and △: point error at 243 mgS/L). **a** Determining $\hat{q}$ and K_I; **b** determining K_S

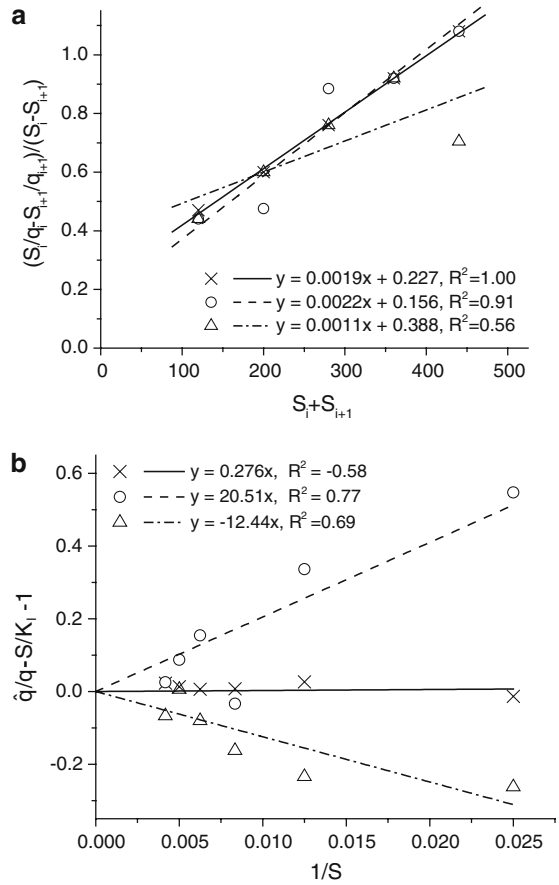

Fig. 2 Influence of one point error with arithmetic-interval data using the K_S-elimination method. (×: point error at 40 mgS/L, ○: point error at 120 mgS/L and △: point error at 240 mgS/L). **a** Determining $\hat{q}$ and K_I; **b** determining K_S

K_S values were even negative in some cases. The sensitivity to one point error probably was caused by the fact that the arithmetic intervals had no values at low substrate concentration (e.g., S value lower than K_S), where q is most sensitive to S. This indicates that an experimental design with arithmetic intervals may have a high risk of generating unreasonable parameter estimates. In addition, a data set that gives very different parameter estimates among the estimation methods might be considered unreliable.

K_S-elimination and K_I-elimination gave similar results when the data were taken with geometric substrate-concentration intervals. Comparing DEM and NLSR, the latter gave better results, particularly when the point error was imposed on the lowest q. This difference probably occurred because the linearized curves expressed in Eqs. 9 and 12, in which q is taken as a reciprocal, amplify the impacts of low q values. An error in q near the middle range of substrate concentration affected DEM estimation the least.

In the case of geometric intervals of S, the estimated S_{max} values were close to each other among the methods (and to the true value) in spite of somewhat significant errors in estimating K_S and K_I. The estimates for K_S and K_I tended to move away from the true values in opposite directions (i.e., when the estimated K_S is smaller than the true value, the estimated K_I is larger than its true value), preventing the estimate of S_{max} from deviating too much (recall that $S_{max} = \sqrt{K_S K_I}$). This pattern is advantageous in a process design for which the goal is to maximize the reaction rate by staying near $S = S_{max}$.

The estimates of q_{max} (q at S_{max}) by the two DEM options were identical, but slightly different from the

Table 3 Summary of test with one 10%-point error

Location of error	Method	Geometric interval of S					Arithmetic interval of S			
		$\hat{q}$	K_S	K_I	S_{max}	q_{max}	$\hat{q}$	K_S	K_I	S_{max}
q For lowest S	K_I elimination	4.3	7.6	126.6	31.0	2.9	4.4	−0.7	119.0	–
	K_S elimination	4.3	7.6	122.4	30.6	2.9	4.4	0.28	119.5	5.8
	NLSR	4.9	9.7	101.9	31.4	3.0	4.5	$1.3*10^{-9}$	114.9	0.0
q For middle S (3rd)	K_I elimination	5.4	10.8	90.1	31.1	3.2	6.5	23.9	70.4	41.0
	K_S elimination	5.4	10.8	88.9	31.0	3.2	6.4	20.5	71.1	38.2
	NLSR	5.0	9.1	97.5	29.8	3.1	6.2	21.6	75.8	40.5
q For highest S	K_I elimination	4.8	9.5	123.5	34.2	3.1	2.7	−10.0	322.6	–
	K_S elimination	4.8	9.5	123.7	34.2	3.1	2.6	−12.4	352.5	–
	NLSR	4.7	9.0	120.8	33.0	3.0	4.1	2.8	142.3	20.0
True value		5	10	100	31.6	3.1	5	10	100	31.6

Units: $\hat{q}$, q_{max} (q at S_{max}) = mgS/mgVSS-d; K_S, K_I, and S_{max} = mgS/L

estimates by NLSR. Though the $\hat{q}$ values estimated by the DEM had slight errors, the $\hat{q}/K_S$ ratios were very close to the true value (0.5). This resulted from the pattern that the estimates for $\hat{q}$ and K_S usually deviated in the same direction from their true values. Having an accurate and stable estimate of $\hat{q}/K_S$ is advantageous when the process goal is to maintain a low substrate concentration, since Eq. 1 simplifies to $q = S\hat{q}/K_S$ when S is very small. Finally, the $\hat{q}K_I$ value tended to be close to the true value (500), which is advantageous for accurate predictions of the reaction rate at a very high substrate concentration, as Eq. 1 simplifies to $q = \hat{q}K_I/S$ at this condition.

With random errors

Table 4 summarizes the results obtained with random errors. Each value represents the mean ± standard deviation from the estimates for the 15 data sets. The two elimination options (K_S- and K_I-elimination) show almost the same results as long as the data intervals for S were geometric. Again, none of the methods gave satisfactory results with arithmetic intervals for S, and the estimates were very different from each other, even between the two DEM options. This observation leads to the conclusion that a significant disagreement in the parameter estimation between the two DEM options indicates significant structural errors in the dataset.

With geometric intervals, the DEM gave as good results as the NLSR did for $\hat{q}$, K_S, and K_I. However, the standard deviations by the DEM were slightly larger, reflecting that the DEM can be more sensitive to random errors: The error in the K_S estimates by the DEM was as high as 40%, whereas that by the NLSR was only 26% (data not shown). However, the average q_{max} values in Table 4 were the same and almost exact, the standard deviations by the DEM

Table 4 Summary of random-error tests

Method	Geometric interval of S					Arithmetic interval of S			
	$\hat{q}$	K_S	K_I	S_{max}	q_{max}	$\hat{q}$	K_S	K_I	S_{max}
K_I elimination	5.2 ± 0.5	10.4 ± 1.6	99.0 ± 16.4	31.7 ± 1.1	3.13 ± 0.13	11.2 ± 21.5	65.8 ± 195.1	106.8 ± 61.5	83.8
K_S elimination	5.2 ± 0.5	10.4 ± 1.6	99.2 ± 15.8	31.7 ± 1.0	3.13 ± 0.13	−1.4 ± 27.7	−52.1 ± 257.5	116.4 ± 74.8	–
NLSR	5.2 ± 0.4	10.3 ± 1.5	96.5 ± 13.7	31.3 ± 1.3	3.13 ± 0.10	5.8 ± 2.0	17.0 ± 18.0	98.7 ± 41.2	41.0
True value	5	10	100	31.6	3.1	5	10	100	31.6

Units: $\hat{q}$, q_{max} = mgS/mgVSS-d; K_S, K_I, and S_{max} = mgS/L

were slightly bigger than that by NLSR. Despite the slight inferiority of the DEM in q_{max} estimation, it gave superior estimates for S_{max}, as seen by mean and standard deviation for S_{max} in Table 4.

In general, the DEM gave good results when the percent errors in q at both extremes of the dataset were small (e.g., set I in Table 2), but relatively poor results when the errors in q were simultaneously large at both extreme (e.g., set II in Table 2). NLSR was not particularly affected by these conditions.

Figure 3 shows the fitting of the data in Table 2 with the parameters estimated by the different methods. Table 5 summarizes the estimated parameters from the two example data sets. With error set I (Fig. 3a), the DEM parameters resulted in a small under-estimation of the experimental data, but the parameter values were closer to the true value than by NLSR. With error set II, which had negative errors in q at the both extremes of substrate concentration, the two DEM options did not fit the data as well as the NLSR method, especially near the maximum q, and the parameter estimates were not as close to the true values.

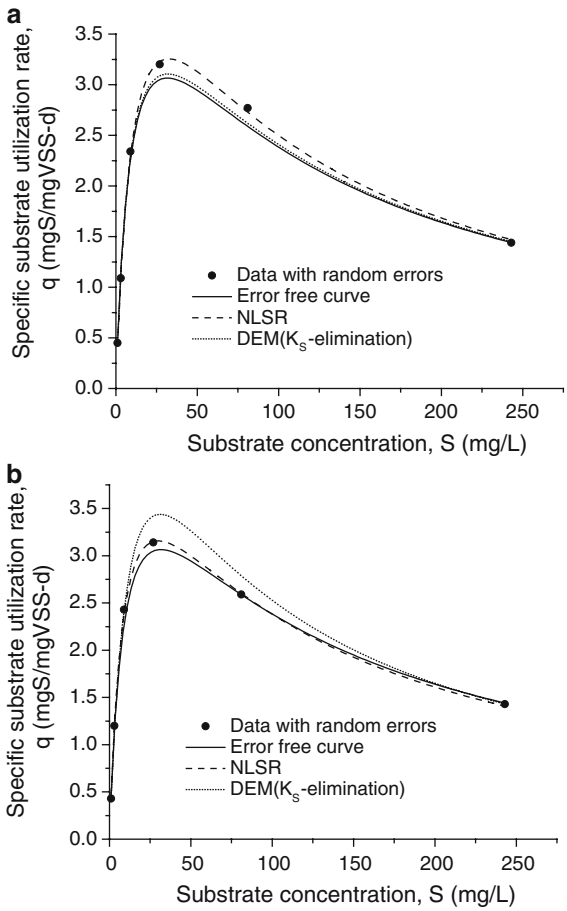

Fig. 3 Best-fit curves by DEM and NLSR for the geometric-interval data sets with random errors in Table 2. **a** Error set I, **b** Error set II

Table 6 summarizes the mean and standard deviation of the squares of residuals of q, $(q_{data} - q_{model})^2$ calculated from the 15 data sets for each substrate concentration. The squares of residuals with the DEM were smaller at the lowest and highest substrate concentrations than were those of NLSR. However, they were substantially larger at the middle range, causing the residual sum of the squares (RSS, $\sum (q_{data} - q_{model})^2$) to be substantially larger than for NLSR. From these observations and those for the one-point error, we conclude that the DEM is most accurate for fitting the data near both extremes (e.g., smallest and largest S), but relatively weak in estimating the $\hat{q}$ when the data contain errors.

Effects of sampling intervals and distribution

Sampling using arithmetic intervals for S did not give satisfactory results when the data contained errors. A critical weakness in the arithmetic-interval data was that it contained few data points at relatively low S (e.g., lower than S_{max}). Thus, we generated two sets of data, one for arithmetic and the other for geometric, for a similar range of S: arithmetic intervals—27, 262, 497 and 732 mg/L, and geometric intervals—27, 81, 243, 729 mg/L. We then imposed six sets of random errors that ranged from 0 to 6% on the q values. The results, summarized in Table 7, make it obvious that the geometric intervals always gave better results than the arithmetic interval did for all methods. Geometric intervals produced always-realistic values, even though the number of samples was small. On the other hand, arithmetic intervals allowed negative values for K_S, and which range for all parameters.

Effects of the error level: comparison DEM options with NLSR

Table 8 shows the parameter estimates by DEM with the 1,000 data sets of three different levels of random errors. With 5% error level, the results were

Table 5 Parameter estimation for the two example data sets in Table 2

Method	Error set I					Error set II				
	$\hat{q}$	K_S	K_I	S_{max}	q_{max}	$\hat{q}$	K_S	K_I	S_{max}	q_{max}
K_I elimination	5.2	10.5	97.1	31.9	3.1	6.5	14.1	68.0	31.0	3.4
K_S elimination	5.2	10.5	96.9	31.9	3.1	6.5	14.0	69.8	31.0	3.4
NLSR	5.7	12.4	85.3	32.5	3.3	5.2	9.9	90.8	29.9	3.2
True value	5	10	100	31.6	3.1	5	10	100	31.6	3.1

Units: $\hat{q}$, q_{max} = mgS/mgVSS-d; K_S, K_I, and S_{max} = mgS/L

Table 6 Average square of residuals for q, calculated from the 15 data sets ($\times 10^{-4}$)

Corresponding S (mg/L)	K_I elimination	K_S elimination	NLSR
1	0 ± 0	0 ± 0	7 ± 8
3	19 ± 22	16 ± 19	29 ± 32
9	106 ± 110	111 ± 119	46 ± 48
27	156 ± 204	177 ± 238	26 ± 28
81	120 ± 139	138 ± 167	31 ± 42
243	1 ± 1	3 ± 2	21 ± 34
Sum (RSS)	403 ± 276	446 ± 337	160 ± 120

Table 7 Summary of estimated parameters with six sets of data with various random errors

Method	Geometric interval of S (27, 81, 243 and 729 mg/L)			Arithmetic interval of S (27, 262, 497 and 732 mg/L)		
	$\hat{q}$	K_S	K_I	$\hat{q}$	K_S	K_I
K_I elimination	4.3–5.2	6.0–10.4	92.6–125.0	3.2–6.6	−1.5–18.3	70.9–175.4
K_S elimination	4.3–5.2	5.6–9.5	91.6–128.7	2.9–6.5	−4.5–18.0	73.7–202.8
NLSR	4.4–7.3	5.7–25.0	59.0–124.9	3.7–16.3	0.7–89.3	26.3–150.6
True value	5	10	100	5	10	100

Table 8 Parameter estimates by DEM with 1,000 data sets (Geometrically spaced data)

Error level	K_I elimination			K_S elimination		
	$\hat{q}$	K_S	K_I	$\hat{q}$	K_S	K_I
5% (σ)	5.1 ± 0.7	10.2 ± 1.9	104.0 ± 9.0	5.1 ± 0.7	10.0 ± 0.9	101.1 ± 20.7
10% (σ)	5.5 ± 1.9	11.6 ± 5.6	108.4 ± 19.5	5.5 ± 1.9	11.3 ± 2.1	106.3 ± 46.3
15% (σ)	7.0 ± 10.1	16.8 ± 31.5	132.6 ± 70.3	7.0 ± 10.1	23.4 ± 7.8	132.8 ± 107.5

satisfactory and similar to those in Table 4. As the error level increased, the quality of estimation got worse in degree, particularly when $\sigma = 15\%$.

Table 9 shows the comparison between DEM and NLSR for the 15% error level. For this analysis, we randomly picked 100 sets of data from the 1,000 sets. The reduction of the sample size was necessary, because the NLSR calculations excessively time-consuming. (This is one reason why DEM is beneficial.) Overall, NLSR was superior to DEM, but the difference was relatively small.

We identified the extreme cases of parameter estimation out of the 100 data sets: 10 cases of inferior DEM estimates and 8 cases of inferior NLSR estimates. The results are given in Table 10. Each of the 10 cases giving unacceptable results with DEM

Table 9 Comparison of DEM and NLSR with 100 data sets of 15% error level

15% (σ)	K_I elimination			NLSR			K_S elimination		
	$\hat{q}$	K_S	K_I	$\hat{q}$	K_S	K_I	$\hat{q}$	K_S	K_I
Average	8.2	21.4	118.5	6.7	16.1	106.6	8.2	14.8	109.7
Standard deviation	23.4	71.5	32.0	4.8	16.9	75.1	23.4	4.1	71.4

Table 10 Extreme estimates by DEM options and NLSR

15% (σ)	DEM: 10 among 100 sets			NLSR: 8 among 100 sets		
	NLSR	K_S elimination	K_I elimination	NLSR	K_S elimination	K_I elimination
$\hat{q}$	6.2	35.8	35.8	23.2	5.5	5.5
K_S	11.9	31.6	121.2	74.9	13.5	12.0
K_I	90.4	8.5	78.7	17.0	95.8	90.1

had a negative $\hat{q}$ or K_I lower than K_S, both of which are unrealistic. The unacceptable results with DEM happened when the data set contained large errors at the both ends of substrate concentrations, since DEM fits those end-points data better (refer to Table 6). However, NLSR gave reasonable estimates when DEM failed. The 8 cases of unacceptable results with NLSR also had $K_I < K_S$. These cases were characterized by errors at specific data points (27 and 81 mg/L of S in our case) within which the apparent maximum value of q was observed. It happened whenever one of the data points had a positive extreme error and the other had a negative extreme error. When NLSR gave unacceptable results, DEM gave reasonable estimates. Thus, the two methods were complementary each other.

As we do not know the true parameter values in reality, DEM can be used as a diagnostic tool to evaluate the appropriateness of the data obtained by comparing the results from NLSR. If both methods give nearly the same results, then the data sets and parameter estimates should be reliable. If the parameter values differ or if one estimation method gives an unrealistic result (e.g., negative $\hat{q}$ or $K_I < K_S$), the data set probably has a serious error that can be diagnosed.

Application of DEM to experimental data

The DEM options were applied to the experimental data in Fig. 4 to estimate the self-inhibition kinetic parameters for nitrification. Evaluating the data according to the linearization equations gave good linearity, as shown in Figs. 5 and 6. The calculated parameter values are presented in Table 11. In particular, the K_I-elimination option gave R^2 values that were almost 1, had no systematic error, and yielded values very close to the estimates by NLSR. Overall, the three methods gave almost identical estimates, implying that the experimental data were reliable. Although the data set had only four data-points, the sampling intervals were geometric and evenly distributed at both sides of S_{max}. The S_{max} estimates indicate that the maximum nitrification rate

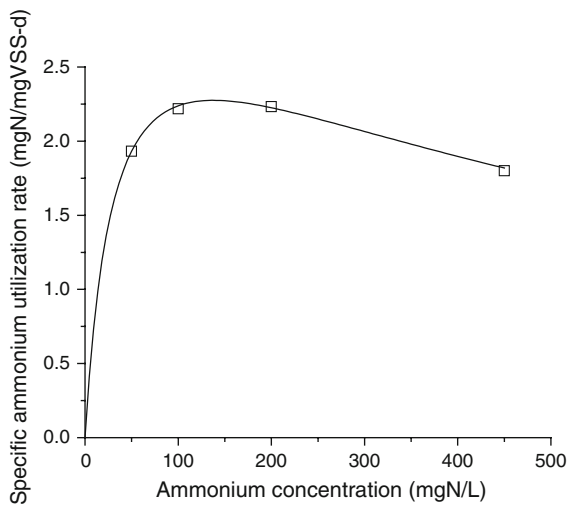

Fig. 4 q versus S for nitrification (Data from Park and Bae (2009)). The *solid line* is a fitting with the parameters estimated by DEM (K_I-elimination)

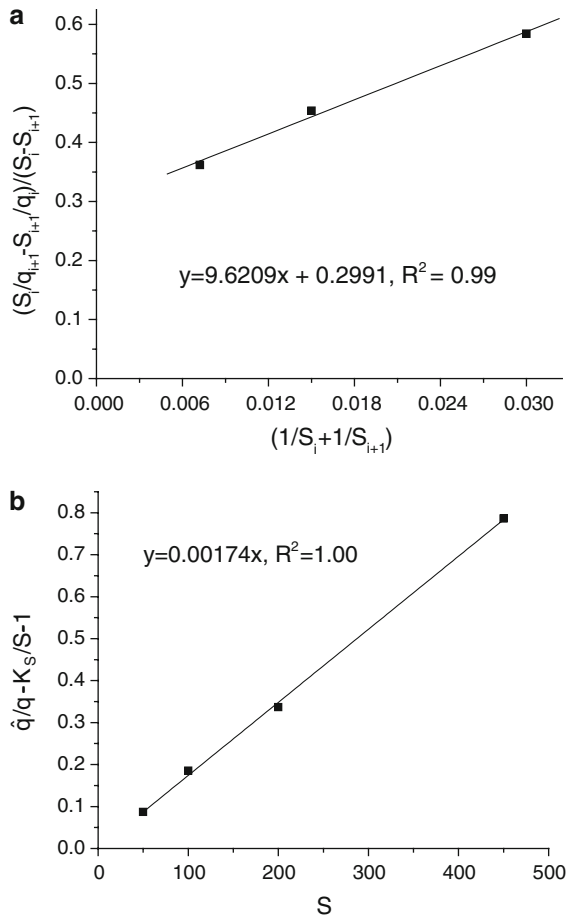

Fig. 5 Linearization of the nitrification data by K_I elimination. **a** Determining $\hat{q}$ and K_S; **b** determining K_I

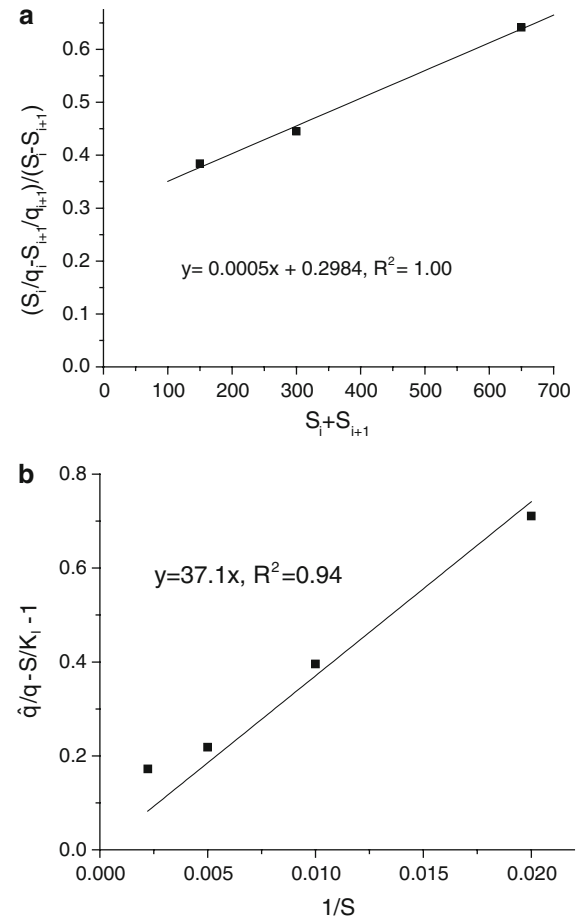

Fig. 6 Linearization of the nitrification data by K_S elimination. **a** Determining $\hat{q}$ and K_I; **b** determining K_S

could be achieved around 140 mgN/L. Finally, the solid line in Fig. 4, drawn using the parameters estimated by the K_I-elimination option, fits the data very well over the entire range.

Figure 7 shows data obtained by Wang et al. (1999) from recombinant cyanobacteria that produced ethylene during photosynthesis. They estimated $\hat{\mu}$ instead of $\hat{q}$ by assuming that ethylene production was proportional to cell growth. The data were linearized almost perfectly ($R^2 \geq 0.98$, no systematic error) by either K_I or K_S elimination. Consequently, the estimated values of parameters were almost identical between the two elimination options and very close to those by NLSR and the estimates by Wang et al. (1999): 2.88 nL/mL/OD730/min for $\hat{\mu}$, 0.516 klux for K_S, 9.29 klux for K_I, and 2.19 klux for S_{max} (Table 11). Consistent with the synthetic data, the S_{max} estimates were more similar than the K_S and K_I values. The solid line in Fig. 7, drawn using the parameters estimated by the K_I-elimination option, gives an excellent fit for the entire range of light intensity.

Data with an arithmetic substrate interval were tested by utilizing the data of phenol degradation by Goudar et al. (2000). The data are shown in Fig. 8. As shown in Figs. 9 and 10, the data after treatment by the DEM did not linearize well. The estimated parameter values by the DEM methods, summarized in Table 11, are quite different from NLSR. Estimated values for $\hat{\mu}$varied as much as 2.4-fold, 3-fold for K_I, and and 23-fold for K_S. Most of all, such a huge difference in results between the methods may indicate that data acquisition was inappropriate and supports that data generation with an arithmetic

Table 11 Parameter values for estimated by all methods

Method	$\hat{q}$	K_S	K_I	S_{max}	q_{max}
Nitrification Park and Bae (2009)					
Elimination method					
K_I	3.34	32.2	588	138	2.28
K_S	3.35	33.0	597	149	2.24
NLSR	3.40	32.7	572	137	2.30
Method	$\hat{\mu}^a$	K_S	K_I	S_{max}	μ^a_{max}
Photosynthetic self-inhibition (Wang et al. 1999)					
Elimination method					
K_I	2.86	0.518	9.44	2.20	1.95
K_S	2.86	0.513	9.46	2.21	1.95
NLSR	2.93	0.541	9.04	2.21	1.97
Phenol degradation (Goudar et al. 2000)					
Elimination method					
K_I	0.43	0.10	0.17	0.13	0.17
K_S	0.57	0.18	0.12	0.15	0.17
NLSR	0.24	0.008	0.37	0.05	0.19

Units for nitrification: $\hat{q}$, q_{max} = mgN/mgVSS-d; K_S, K_I and S_{max} = mgN/L

Units for photosynthesis: $\hat{\mu}$, μ_{max} = nL/mL/OD730/min; K_S, K_I and S_{max} = klux

Units for Phenol degration: $\hat{\mu}$, $\mu_{max} = h^{-1}$; K_S, K_I and S_{max} = g/L

[a] $\hat{\mu}$ and μ_{max} are, respectively, the maximum and observed maximum specific ethylene production rates

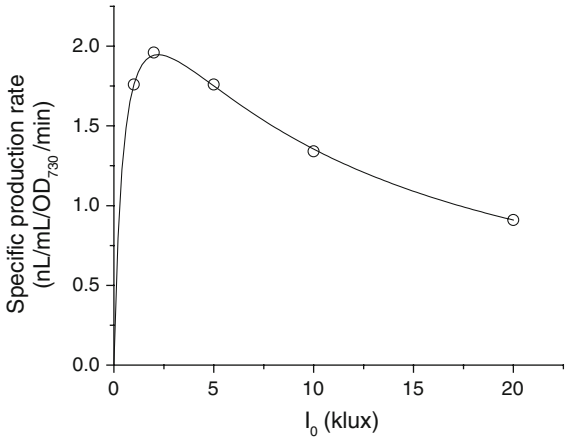

Fig. 7 Specific production rate of ethylene by recombinant cyanobacteria as a function of light intensity (Data from Wang et al. (1999)). The *solid line* is a fitting with the parameters estimated by the DEM (K_I-elimination)

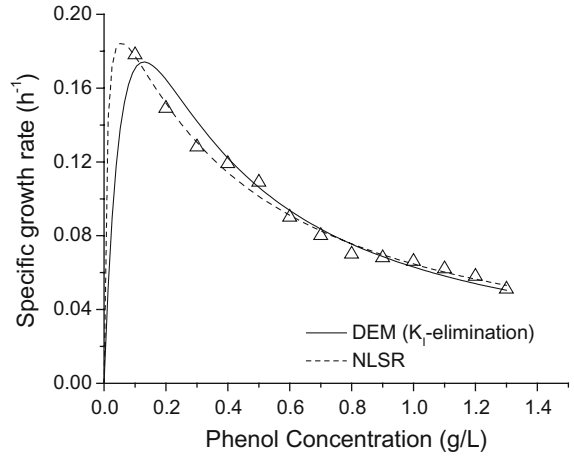

Fig. 8 Microbial specific growth rate at various initial phenol concentrations. (Data from Goudar et al. (2000)). The *solid line* is a fitting with the parameters estimated by the DEM (K_I-elimination). The *dashed line* is a fitting by NLSR

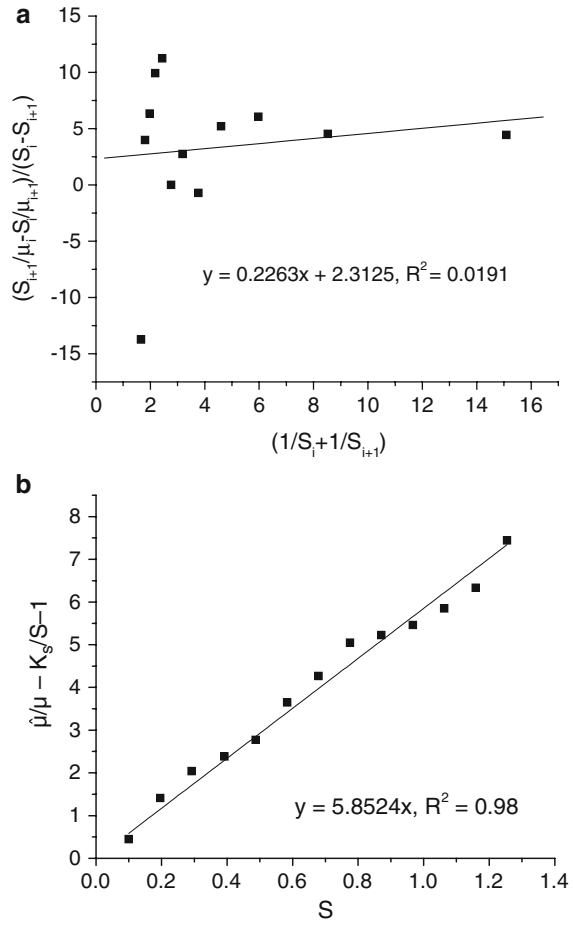

Fig. 9 Linearization of the data for phenol degradation using K_I elimination. **a** Determining $\hat{q}$ and K_S; **b** determining K_I

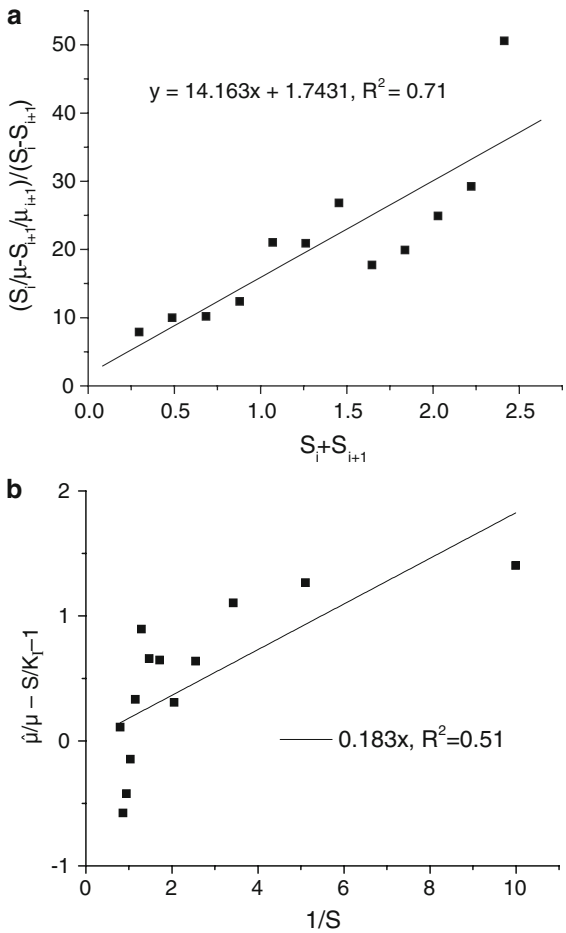

Fig. 10 Linearization of the data on phenol degradation using the K_S elimination. **a** Determining $\hat{q}$ and K_I; **b** determining K_S

interval for substrate concentration can introduce problems when applying the DEM. The solid curve in Fig. 8, which was drawn using the estimates by the K_I-elimination option, fits the data poorly in comparison with that in Figs. 4 or 7. Quite significant deviations occur with high-μ data-points compared to the dashed curve by NLSR, which again indicated that the data sets were inappropriate for parameter estimation in terms of the substrate intervals and/or magnitude of errors.

Conclusions

The differential elimination method (DEM) linearizes the substrate self-inhibition equation so that $\hat{q}$, K_S, and K_I can be estimated without using non-linear least square regression (NLSR). The K_I- and K_S-elimination options estimated the parameter values exactly when the data did not contain any error. When one-point or random errors were not too large, the DEM worked as well as NLSR, as long as the data were acquired with geometric intervals of substrate concentration, and both elimination options gave almost the same results. The DEM was most accurate for fitting the data near both extremes of S, but relatively weaker in estimating the observed maximum substrate utilization rate, q_{max}. The estimates for S_{max}, the concentration at which q_{max} is observed, were relatively invariant among the methods, even when K_S and K_I differed. When the intervals were arithmetic (i.e., equal) and the data contained errors, the DEM options and NLSR estimated the parameters poorly and differently. Thus, collecting data with an arithmetic interval increases the risk of poor parameter estimation. The best estimation was obtained when the substrate intervals were geometric and the data points evenly distributed on both sides of S_{max}. Besides providing a reasonable substitute for NLSR, the DEM options also can be used as a tool to diagnose the quality of experimental data by comparing its estimates between the DEM options, or, more rigorously, to those from NLSR.

Acknowledgments "This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD)." (KRF-2006-352-D00122).

References

Andrews JF (1968) A mathematical model for continuous culture of microorganisms utilizing inhibitory substrates. Biotechnol Bioeng 10(6):707–723

Bailey JE, Ollis DF (1986) Biochemical engineering fundamentals. McGraw-Hill, New york, p 928

Bates DM, Watts DG (1988) Nonlinear regression and its applications. Wiley, New York

Chen SC, Li KH, Fang HY (2005) Growth kinetics of EDTA biodegradation by *Burkholderia cepacia*. World J Microbiol Biotechnol 21(1):11–16

Futamata H, Harayama S, Watanabe K (2001) Diversity in kinetics of trichloroethylene-degrading activities exhibited by phenol-degrading bacteria. Appl Microbiol Biotechnol 55(2):248–253

Goudar CT, Ganji SH, Pujar BG, Strevett KA (2000) Substrate inhibition kinetics of phenol biodegradation. Water Environ Res 72(1):50–55

Haldane JBS (1930) Enzymes. Longmans, London

Kotturi G, Robinson CW, Inniss WE (1991) Phenol degradation by a psychrotrophic strain of *Pseudomonas Putida*. Appl Microbiol Biotechnol 34(4):539–543

Lineweaver H, Burk D (1934) The determination of enzyme dissociation constants. J Am Chem Soc 56:658–666

Park S, Bae W (2009) Modeling kinetics of ammonium oxidation and nitrite oxidation under simultaneous inhibition by free ammonia and free nitrous acid. Process Biochem 44(6):631–640

Vadivelu VM, Keller J, Yuan ZG (2006) Effect of free ammonia and free nitrous acid concentration on the anabolic and catabolic processes of an enriched Nitrosomonas culture. Biotechnol Bioeng 95(5):830–839

Wang JS, Araki T, Ogawa T, Matsuoka M, Fukuda H (1999) A method of graphically analyzing substrate-inhibition kinetics. Biotechnol Bioeng 62(4):402–411