VOL. 13 NO. 9,
AUGUST, 2026
E-ISSN 3026-958X
P-ISSN 3027-1169

African Scholar Publications & Research International
www.africanscholarpub.com

Journal of Science Innovation & Technology Research (JSITR)

Multi-Criteria Assessment of Biochemical Oxygen Demand Kinetic Parameter Estimation: Evidence from Seven Wastewater Assays

Mbeyi, Paschal Agha; & Ajah, Uche Christian
Department of Civil Engineering Technology, Akanu Ibiam Federal Polytechnic Unwana, Ebonyi State, Nigeria
Corresponding Author: paxkalville@yahoo.com
DOI: https://doi.org/10.70382/ajsitr.v13i9.089

Abstract
Biochemical oxygen demand (BOD) kinetics are widely used to describe oxygen consumption during biodegradation of wastewater organic matter and to support wastewater-treatment and receiving-water assessments. This study presents a seven wastewater BOD assays and comparing least-squares, Thomas and nonlinear regression approaches using kinetic parameters and three goodness-of-fit measures: sum of squared errors (SS), coefficient of determination (CD), and model selection criterion (MSC). The waste water came from food-processing effluents with BOD_5 values of 629–938 mg L^{-1}. Across the seven curves, nonlinear regression produced the lowest mean SS (2,423.0), compared with 4,131.1 for least squares and 4,847.2 for Thomas; the highest mean CD (0.99708); and the highest mean MSC (5.411). A Friedman test applied to the paired method comparisons indicated a significant overall difference among methods ($\chi^2 = 10.571$, $p = 0.005$), with Kendall's $W = 0.755$, indicating a large method effect. Exact paired post-hoc Wilcoxon tests showed that nonlinear regression outperformed both least-squares and Thomas methods for all three criteria. The analysis supports nonlinear estimation as the most consistently accurate of the three approaches, while also demonstrating that reliance on a single goodness-of-fit statistic can obscure method differences.

Keywords: Biochemical Oxygen Demand; BOD kinetics; Wastewater; Nonlinear Regression; Thomas method; Least Squares; Model Selection; Goodness of Fit;

Introduction

The modern economy is facing severe water pollution challenges because of population growth, agriculture, and industrial activities. Increasing the efficiency of wastewater treatment plants is recognized as the most significant method of controlling the discharge of pollutants into the environment (Adibimanesh, et al., 2023). Biochemical oxygen demand (BOD) remains an important indicator of biodegradable organic loading in wastewater and is used in wastewater-treatment assessment and water-quality modelling. The conventional BOD_5 test estimates oxygen consumption over a five-day incubation period, but the laboratory measurement is inherently time-consuming. Recent research has therefore continued to examine predictive and modelling approaches that can improve the speed and reliability of wastewater-quality assessment.

The characterization of organic pollution in wastewater is an important component of environmental engineering and wastewater-treatment management. Among the commonly used indicators of organic pollution, biochemical oxygen demand (BOD) provides an estimate of the oxygen required by microorganisms to biologically degrade biodegradable organic matter under specified conditions. BOD is therefore widely used as an indicator of wastewater organic strength and as an important parameter for evaluating wastewater-treatment performance and water-quality conditions. The standard five-day biochemical oxygen demand test (BOD_5) remains one of the established procedures for determining oxygen consumption associated with biodegradable organic matter in wastewater (APHA, AWWA, & WEF, 2017). The importance of BOD_5 in wastewater characterization and treatment has also been demonstrated by recent studies that use BOD_5 as a key indicator for wastewater-treatment plant performance, modelling and process optimization (Qambar & Al Khalidy, 2022; Aghdam et al., 2023).

Recent research has consequently investigated mathematical, statistical and machine-learning approaches for estimating BOD_5 more rapidly than conventional laboratory testing (Qambar & Al Khalidy, 2022; Aghdam et al., 2023). Periodic measurements of dissolved oxygen can therefore be used to describe the BOD–time relationship and develop kinetic models capable of estimating ultimate oxygen demand and the rate at which oxygen is consumed. Cutrera et al. (1999) demonstrated the importance of following the BOD curve over time and noted that experimental conditions such as dilution, toxic substances and microbial acclimatization can influence the determination of BOD kinetic parameters.

The mathematical description of BOD kinetics is commonly based on a first-order reaction model. Under this formulation, the rate of change of the remaining oxygen demand is proportional to the amount of biodegradable material remaining in the system. First-order formulations have long been used to describe BOD exertion and to estimate the kinetic constants required for modelling oxygen consumption in wastewater and receiving-water systems (Thomas, 1950; Cutrera et al., 1999).

Reliable estimation of these parameters is therefore important because errors in either parameter can influence predictions of oxygen consumption and the interpretation of wastewater biodegradation characteristics. The importance of accurately estimating BOD kinetic constants has been recognized in earlier studies comparing different mathematical formulations and parameter-estimation procedures (Thomas, 1950; Moore et al., 1950; Siwiec et al., 2012).

Several approaches have been developed for estimating BOD kinetic parameters. Among the commonly used procedures are linear or least-squares regression, the Thomas method, and nonlinear regression. Linear regression can be applied to transformed forms of the BOD equation, while the Thomas method uses a

specific transformation of the BOD–time relationship to permit graphical or linear estimation of the kinetic parameters (Thomas, 1950). The Thomas approach is based on plotting a transformed quantity involving (t/y)1/3 against time and obtaining the kinetic constants from the slope and intercept of the resulting relationship (Thomas, 1950; Cutrera et al., 1999). Nonlinear regression, in contrast, estimates the kinetic coefficients directly from the original nonlinear BOD equation. Comparative studies have shown that different mathematical formulations can produce different estimates of BOD kinetic constants, demonstrating the importance of the parameter-estimation procedure used (Cutrera et al., 1999; Siwiec et al., 2012).

The choice of parameter-estimation method can therefore have a substantial effect on the resulting values of k and L0. Cutrera et al. (1999) reported marked differences among the kinetic coefficients obtained from least-squares, Thomas and nonlinear regression procedures and emphasized that direct inspection of the estimated kinetic parameters alone was insufficient for determining which method provided the best representation of the experimental BOD data. This observation is important because a method producing a plausible kinetic parameter does not necessarily provide the best representation of the observed BOD curve. Consequently, objective goodness-of-fit criteria are required to evaluate competing estimation procedures.

Several statistical criteria can be employed for this purpose. The sum of squared errors (SS) measures the discrepancy between observed and calculated BOD values, with lower values indicating smaller fitting errors. The coefficient of determination (CD**)** provides an indication of the degree of agreement between observed and calculated values, with values approaching unity generally indicating a stronger fit. The model selection criterion (MSC) provides an additional basis for comparing competing models and is associated with information-theoretic model-selection principles (Akaike, 1974). Cutrera et al. (1999) specifically used total error, coefficient of determination and MSC to compare the performance of the three BOD kinetic parameter-estimation methods and noted the usefulness of MSC as an additional criterion because of its treatment of model complexity and its independence from the numerical scale of the measurements.

An important methodological issue, therefore, is that comparison of BOD kinetic estimation procedures should not depend on a single goodness-of-fit statistic. A method may perform favourably according to one criterion while showing a different ranking under another. The use of several complementary performance measures can provide a more comprehensive basis for identifying the method that gives the most consistent representation of observed BOD kinetics. This is particularly relevant because mathematical modelling and prediction of wastewater quality increasingly support wastewater-treatment plant operation, process optimization and environmental management (Aghdam et al., 2023; Ching et al., 2022).

The present study therefore undertakes a multi-criteria secondary-data assessment of three BOD kinetic parameter-estimation procedures—least-squares regression, the Thomas method and nonlinear regression—using seven published wastewater assays. The study evaluates the methods using SS, CD and MSC and supplements the descriptive comparison with paired non-parametric statistical analysis. The study also examines the resulting estimates of the kinetic parameters k and L0, thereby distinguishing between differences in parameter estimates and differences in overall model-fitting performance.

The principal contribution of this study is the provision of a transparent and reproducible statistical framework for comparing alternative BOD kinetic estimation procedures. Rather than relying exclusively on individual goodness-of-fit values, the study integrates multiple performance criteria and formal statistical testing to determine the consistency of method performance across the seven

wastewater assays. Such a framework is particularly relevant in contemporary wastewater modelling, where accurate prediction and characterization of BOD remain important for efficient wastewater-treatment operation and environmental management (Qambar & Al Khalidy, 2022; Aghdam et al., 2023). The findings are expected to contribute to more rigorous selection of BOD kinetic estimation procedures for wastewater characterization and environmental engineering applications

Aim

The aim is to provide a reproducible statistical reassessment of three BOD kinetic parameter-estimation methods using the seven wastewater assays reported in the source study.

- compare least-squares, Thomas and nonlinear regression using SS, CD and MSC;
- quantify average differences in goodness-of-fit between nonlinear regression and the alternative methods;
- test whether the paired method differences are statistically significant;
- compare the resulting estimates of k and L_o; and
- identify implications and limitations for wastewater kinetic modelling and future studies.

Specific objectives

The study seeks to:

1. determine and compare the k and L0 values obtained using the three estimation methods;
2. compare the methods using sum of squared errors (SS), coefficient of determination (CD), and model selection criterion (MSC);
3. establish the ranking of the three methods across the seven wastewater assays;
4. statistically determine whether significant differences exist among the estimation methods; and
5. identify the most consistent parameter-estimation procedure for the wastewater assays considered.

Theoretical Basis of BOD Kinetics

The study by Cutrera et al. (1999) provides a useful dataset for such an assessment. Their investigation considered BOD analyses of food-processing effluents from bakeries and fish-processing plants, with reported BOD_5 concentrations ranging from 629 to 938 mg/L. The study compared least-squares, Thomas and nonlinear regression procedures and evaluated their performance using total error, coefficient of determination and model selection criterion. The published results showed that nonlinear regression produced the smallest error, highest CD and highest MSC for each of the seven experimental curves.

For kinetic interpretation, the classical first-order BOD formulation describes oxygen uptake as an asymptotic process controlled by a rate coefficient and ultimate oxygen demand. The integrated model can be written as $y = L_o(1 - \exp(-kt))$, where y is oxygen consumed at time t, k is the reaction-rate constant and L_o is the ultimate BOD. This formulation remains useful for engineering interpretation because k characterizes the apparent rate of biodegradation while L_o represents the ultimate oxygen requirement.

The BOD process was represented in the source study using a first-order kinetic relationship. The differential equation is expressed as:

$\frac{dL}{dt} = -kL$

where L represents the remaining oxygen demand at time t, and k is the reaction-rate constant. Integration of the equation gives the conventional BOD relationship:

$y=L0(1-e-kt)$

where:

- y = oxygen consumed at time t;
- L0 = ultimate biochemical oxygen demand;
- k = BOD reaction-rate constant; and
- t = elapsed time.

This first-order formulation was the basis for estimating the kinetic parameters examined in the present reanalysis.

The parameter k represents the apparent rate at which biodegradable organic matter is oxidized, while L0 represents the ultimate oxygen demand associated with the biodegradable material.

BOD Kinetic Parameter-Estimation

Three parameter-estimation approaches reported in the source study were considered.

Least-Squares Regression

The least-squares approach estimates the kinetic coefficients by fitting an appropriate linearized form of the BOD equation. The method determines the parameter values that minimize the discrepancy between observed and calculated values.

The source study identifies linear regression as one of the commonly used approaches for determining k and L0.

Thomas Method

The Thomas method is based on a transformation of the BOD relationship. According to the source paper, the transformation involves plotting:

$(yt)1/3$

against t, after which the resulting relationship is fitted as a straight line. The intercept and slope of the fitted line are subsequently used to estimate the BOD kinetic parameters.

The Thomas method is therefore mathematically different from direct nonlinear fitting because it relies on transformation of the original BOD relationship before parameter estimation.

Nonlinear Regression

Nonlinear regression estimates the kinetic parameters directly from the original BOD equation rather than first transforming it into a linear form.

The source study reports that nonlinear regression was performed using an optimization program, while the linear and Thomas calculations were implemented using an electronic spreadsheet.

The nonlinear approach is particularly relevant to this investigation because it allows the fitted parameters to be obtained directly from the original nonlinear relationship between BOD and time.

Goodness-of-Fit Criteria

To provide a multi-criteria assessment, three goodness-of-fit measures reported in the source study were retained.

Sum of Squared Errors

The sum of squared errors was calculated conceptually as:

$$SS=i=1\sum n(Yobs,i-Ycal,i)2$$

where Yobs,i represents an observed BOD value and Ycal,i represents the corresponding calculated value.

A lower Sum of Squared (SS) indicates a smaller discrepancy between observed and calculated BOD values and therefore better fitting performance.

The source study refers to this measure as the total error and defines it as the sum of squares of the differences between predicted and experimental values.

Coefficient of Determination

The coefficient of determination was used as a second measure of model performance:

$$CD=1-\sum(Yobs-\bar{Y}obs)2\sum(Yobs-Ycal)2$$

where $\bar{Y}obs$ is the mean observed BOD value.

Values closer to 1 indicate stronger agreement between observed and calculated values.

Model Selection Criterion

The model selection criterion (MSC) was used as the third performance measure. The source study adopted MSC following Akaike's information-theoretic approach. Unlike SS and CD, MSC incorporates a penalty associated with the number of parameters and is not dependent on the numerical scale of the measurements.

For interpretation:

lower SS = better performance; higher CD = better performance; and higher MSC = better performance.

Using the three indicators simultaneously allows the estimation procedures to be assessed from complementary perspectives rather than relying on a single goodness-of-fit measure.

Lower SS indicates smaller fitting error, whereas higher CD and MSC indicate better fit. MSC additionally penalizes model complexity and is therefore useful as a model-selection statistic.

Data source and study design

This study is a secondary-data reanalysis. The dataset was transcribed from Tables 1 and 2 of Cutrera et al. (1999), published in Water SA. The source study analysed food-processing effluents from bakeries and fish-processing plants, with BOD_5 values reported between 629 and 938 mg L^{-1}. BOD analysis was performed according to the then-current Standard Methods procedure and dissolved oxygen was measured using the azide modification of the Winkler method. The source paper reported seven BOD curves and estimates of k and L_0 for each of the three estimation methods.

Only numerical values explicitly available in the uploaded article were used. The present analysis does not reconstruct or invent the underlying raw time-series observations. Consequently, no new laboratory testing, imputation of missing raw BOD observations, or independent recalculation of the original nonlinear fits is claimed.

Research Design

This study adopted a secondary-data assessment/reanalysis design. The analysis was based exclusively on previously published biochemical oxygen demand (BOD) kinetic data reported by Cutrera, Manfredi, del Valle and González in *Water SA* (1999). The study did not involve the collection of new wastewater samples, laboratory experimentation or generation of new BOD measurements. Instead, the published results were systematically extracted and statistically reassessed to provide a multi-criteria comparison of alternative BOD kinetic parameter-estimation procedures.

The use of a secondary-data assessment/reanalysis approach was considered appropriate because the source study provides a directly comparable set of seven BOD assays evaluated using three parameter-estimation procedures. These procedures were least-squares regression, the Thomas method and nonlinear regression. The source study also reported three performance indicators—total error/sum of squared errors (SS), coefficient of determination (CD), and model selection criterion (MSC)—which provide a suitable basis for comparative statistical analysis.

All experimental information and numerical results used in the present analysis are treated as published secondary data, and the contribution of this study lies in their systematic multi-criteria and statistical reassessment.

Source of Data

The dataset was obtained from the published article by Cutrera et al. (1999), entitled **"**On the determination of the kinetic parameters for the BOD test**,"** published in *Water SA*, Volume 25, Number 3.

The original investigation analysed BOD characteristics of food-processing effluents from and fish-processing plants. The reported BOD_5 concentrations ranged from 629 to 938 mg/L. The original samples were analysed using procedures based on Standard Methods, while dissolved oxygen was measured using the azide modification of the Winkler method.

The source article presented seven experimental BOD curves, designated Curves 1–7. For each curve, the original investigators estimated the kinetic parameters k and L0 using the three alternative methods and subsequently calculated the corresponding goodness-of-fit statistics.

The present study extracted the following variables from the published tables:

- BOD kinetic rate constant, k;
- ultimate BOD, L0;
- sum of squared errors, SS;
- coefficient of determination, CD; and
- model selection criterion, MSC.

BOD kinetic model

The source study used the first-order BOD formulation. In differential form, the model assumes that the rate of oxygen uptake is proportional to the remaining oxygen demand. The integrated form is:

$y(t) = L_o [1 - \exp(-kt)]$

where y(t) is oxygen consumed at time t, L_o is ultimate oxygen demand and k is the apparent reaction-rate constant.

Estimation methods

Three estimation procedures were compared: (i) least-squares fitting of a transformed BOD expression; (ii) the Thomas method, which uses a transformed relationship between (t/y)^(1/3) and t; and (iii)

nonlinear regression, in which the kinetic parameters are estimated directly by minimizing the squared discrepancy between observed and calculated BOD values.

Statistical analysis

The seven curves were treated as matched observations because every curve was evaluated using all three estimation methods. For each goodness-of-fit metric, arithmetic means and standard deviations were calculated. Method rankings were determined within each curve. A Friedman test was used to test the global null hypothesis of equal method performance because the sample consisted of seven matched curves and normality of the summary differences cannot be established from the published data. Kendall's coefficient of concordance (W) was used as an effect-size measure. Exact paired Wilcoxon signed-rank tests were used for pairwise comparisons, with Bonferroni-adjusted significance level $\alpha = 0.0167$. Percentage improvement of nonlinear regression relative to each alternative was calculated as (alternative − nonlinear)/alternative × 100 for SS and as (nonlinear − alternative)/alternative × 100 for CD and MSC.

Results

The study results were obtained using a transparent statistical framework. Rather than reproducing the original experiment, it treats the seven reported curves as paired benchmark cases and evaluates whether the apparent superiority of nonlinear regression remains statistically supported when the three published goodness-of-fit criteria are analysed jointly. This approach is relevant because current wastewater modelling increasingly emphasizes rigorous model comparison, predictive reliability and transparent statistical evaluation.

kinetic parameter estimates

Table 1. Mean and standard deviation of kinetic parameters across the seven published BOD curves

Parameter	Least squares	Thomas	Non-linear
Mean k (day^{-1})	0.4811	0.4214	0.4722
SD of k	0.1016	0.0882	0.0923
Mean L_o ($mg\ L^{-1}$)	956.86	1003.72	942.91
SD of L_o ($mg\ L^{-1}$)	114.76	107.39	110.65

The mean k estimates were 0.4811 day^{-1} for least squares, 0.4214 day^{-1} for Thomas and 0.4722 day^{-1} for nonlinear regression. The corresponding mean L_o estimates were 956.86, 1003.72 and 942.91 mg L^{-1}, respectively. Thus, nonlinear regression produced k values close to the least-squares estimates on average but consistently higher than the Thomas estimates, while its mean L_o was lower than both alternatives.

Goodness-of-fit performance

Table 2. Descriptive comparison of goodness-of-fit statistics across the seven curves.

Metric	Least squares	Thomas	Non-linear
Mean SS	4,131.1	4,847.2	2,423.0
SD of SS	4,646.25	5,426.90	2,777.11
Mean CD	0.99509	0.99435	0.99708
SD of CD	0.00232	0.00236	0.00203

Metric	Least squares	Thomas	Non-linear
Mean MSC	4.751	4.529	5.411
SD of MSC	0.507	0.694	0.793

Nonlinear regression had the lowest SS for every curve, the highest CD for every curve and the highest MSC for every curve. The mean SS was 2,423.0 for nonlinear regression compared with 4,131.1 for least squares and 4,847.2 for Thomas. Mean CD values were 0.99708, 0.99509 and 0.99435 for nonlinear, least-squares and Thomas methods, respectively. Mean MSC values were 5.411, 4.751 and 4.529, respectively.

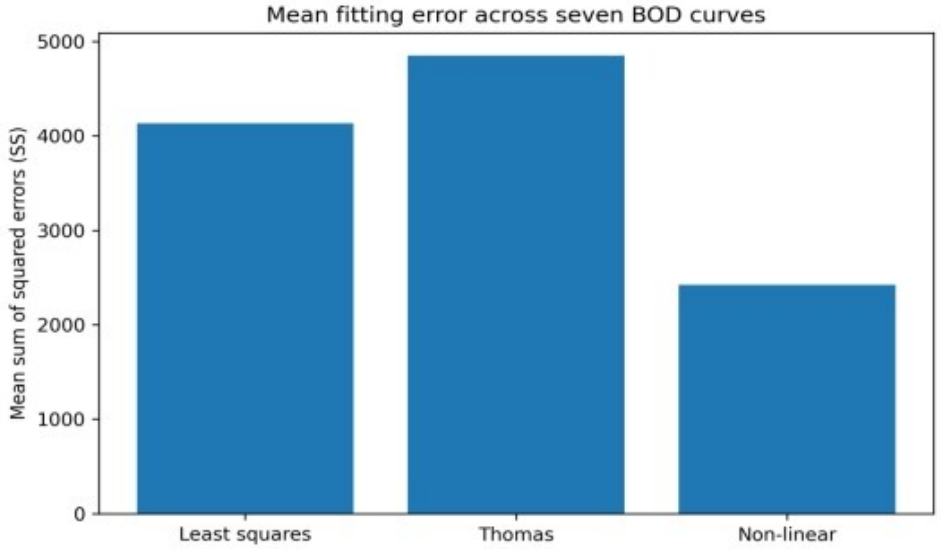


Figure 1. Mean sum of squared errors across the seven BOD curves.

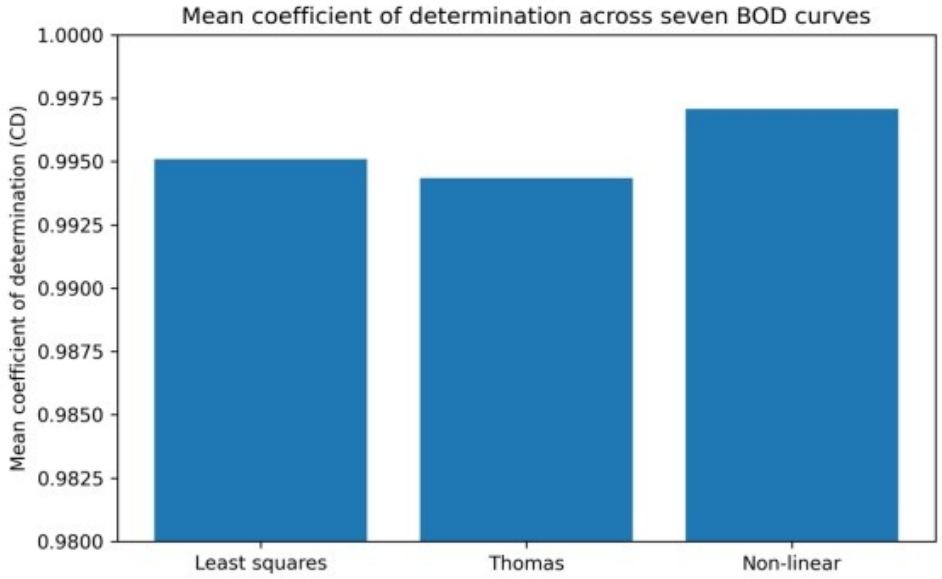


Figure 2. Mean coefficient of determination across the seven BOD curves.

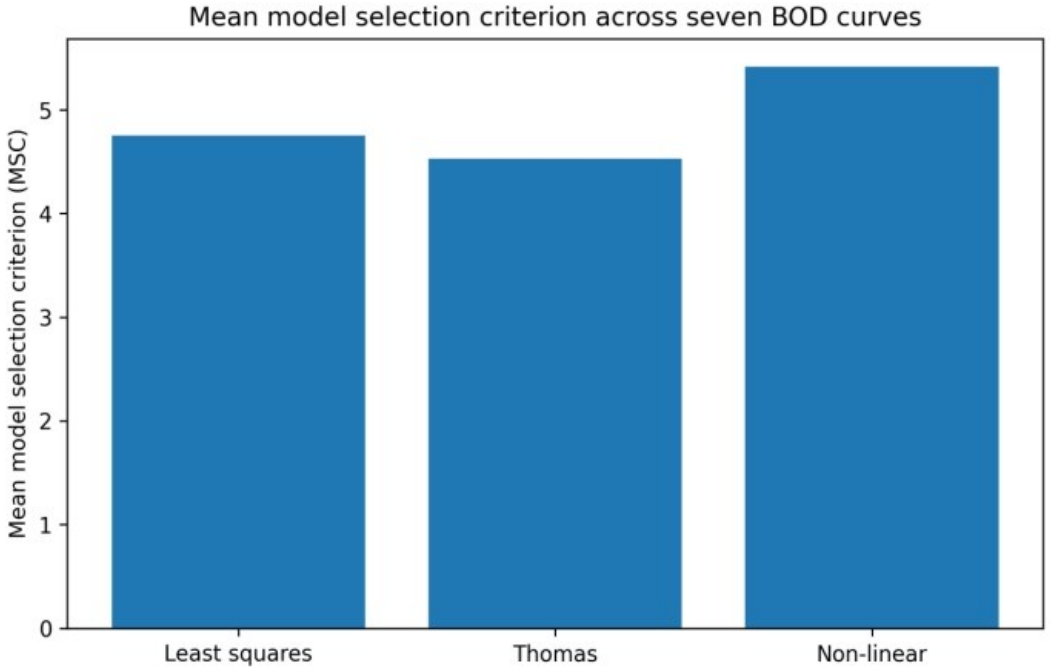


Figure 3. Mean model selection criterion across the seven BOD curves.

Statistical significance of method differences

Table 3. Exact paired Wilcoxon p-values for the published goodness-of-fit metrics.

Comparison	SS p	CD p	MSC p
Least squares vs Thomas	0.2188	0.2188	0.2969
Least squares vs Non-linear	0.0156	0.0156	0.0156
Thomas vs Non-linear	0.0156	0.0156	0.0156

Inferential Statistical Analysis

The seven wastewater assays were treated as matched observations because each assay was evaluated using all three methods.

A Friedman test was therefore applied separately to SS, CD and MSC.

Table 4. Friedman test results

Criterion	Friedman $\chi 2$	df	p-value	Kendall's W	Decision
SS	10.571	2	**0.0051**	**0.755**	Significant
CD	10.571	2	**0.0051**	**0.755**	Significant
MSC	10.571	2	**0.0051**	**0.755**	Significant

The Friedman test produced:

$\chi 2(2)=10.571, p=0.0051$

for each of the three criteria.

Since:
$p<0.05$
the null hypothesis of equivalent performance among the three methods is rejected.
The corresponding Kendall's coefficient of concordance was: $W=0.755$
indicating a strong degree of agreement in method ranking across the seven assays.

The Friedman test gave $\chi^2(2) = 10.571$ and $p = 0.005$ for each of the three goodness-of-fit metrics because the within-curve ranking was identical across SS, CD and MSC. Kendall's W was 0.755, indicating strong agreement in the method ranking and a large method effect. The exact paired Wilcoxon comparison between nonlinear and least-squares methods yielded $p = 0.0156$ for each metric; the comparison between nonlinear and Thomas also yielded $p = 0.0156$ for each metric. After Bonferroni correction for three pairwise comparisons, the adjusted threshold is 0.0167, so both nonlinear-versus-least-squares and nonlinear-versus-Thomas differences remain statistically significant at the 5% family-wise level. The least-squares versus Thomas comparisons were not significant ($p = 0.219$ for SS and CD; $p = 0.297$ for MSC).

Relative improvement of nonlinear regression

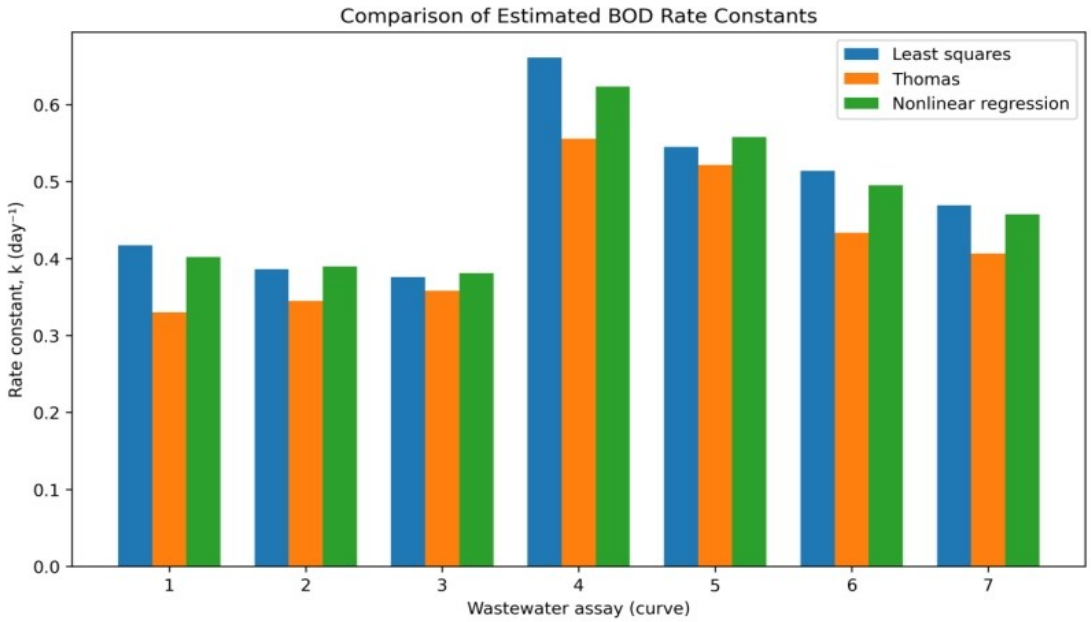
Relative to least squares, nonlinear regression reduced SS by an average of 43.94% and increased CD and MSC by 0.200% and 13.67%, respectively. Relative to Thomas, nonlinear regression reduced SS by an average of 48.50% and increased CD and MSC by 0.275% and 20.42%, respectively. The absolute improvement in CD appears numerically small because all three methods already produced CD values close to unity; MSC and SS therefore provide additional discrimination among the methods.

Estimated BOD Kinetic Parameters

Table 5. BOD kinetic parameters (values of k and L0) obtained from the three estimation methods

Curve	k Least Squares	k Thomas	k Nonlinear	L0 Least Squares	L0 Thomas	L0 Nonlinear
1	0.4170	0.3299	0.40161	753.611	825.50	748.04
2	0.3860	0.3450	0.38959	967.881	1,009.22	946.19
3	0.3760	0.3579	0.38089	1,023.712	1,042.98	1,003.40
4	0.6610	0.5556	0.62346	849.375	902.53	845.23
5	0.5450	0.5215	0.55783	993.819	1,017.06	968.23
6	0.5140	0.4332	0.49477	1,032.315	1,092.53	1,023.00
7	0.4690	0.4065	0.45732	1,077.287	1,136.21	1,066.30
Mean	**0.4811**	**0.4214**	**0.4722**	**956.86**	**1,003.72**	**942.91**

Source: Secondary-data extraction from Cutrera et al. (1999).



Figre 4; Comparison of Estimated BOD rate constants

The results demonstrate noticeable differences in the estimated kinetic parameters depending on the estimation procedure. The Thomas method generally produced lower estimates of k and higher estimates of L0, whereas nonlinear regression produced values that were generally intermediate for k and lower for L0.

The mean rate constants were 0.4811 day^{-1}, 0.4214 day^{-1}, and 0.4722 day^{-1} for least squares, Thomas and nonlinear regression, respectively. The corresponding mean ultimate BOD values were 956.86, 1,003.72 and 942.91 mg/L.

This demonstrates that the choice of estimation method can influence the numerical values of the kinetic parameters even when the same BOD assay is analysed.

Thus, nonlinear regression reduced the average squared fitting error by approximately 41.35% relative to least squares and 50.01% relative to Thomas.

Comparison of Sum of Squared Errors

The published SS values for the seven assays are presented in Table 2.

Table 6. Sum of squared errors for the three estimation methods

Curve	Least Squares	Thomas	Nonlinear Regression
1	2,808.8	2,962.5	2,005.2
2	2,851.8	2,792.2	2,160.1
3	958.4	850.3	513.5
4	1,714.9	3,069.5	337.2
5	2,260.2	3,472.5	1,293.6
6	3,864.9	3,818.3	2,171.0
7	14,459.0	16,965.3	8,480.6
Mean	4,131.14	4,847.23	2,423.03

Source: Cutrera et al. (1999), Table 2.

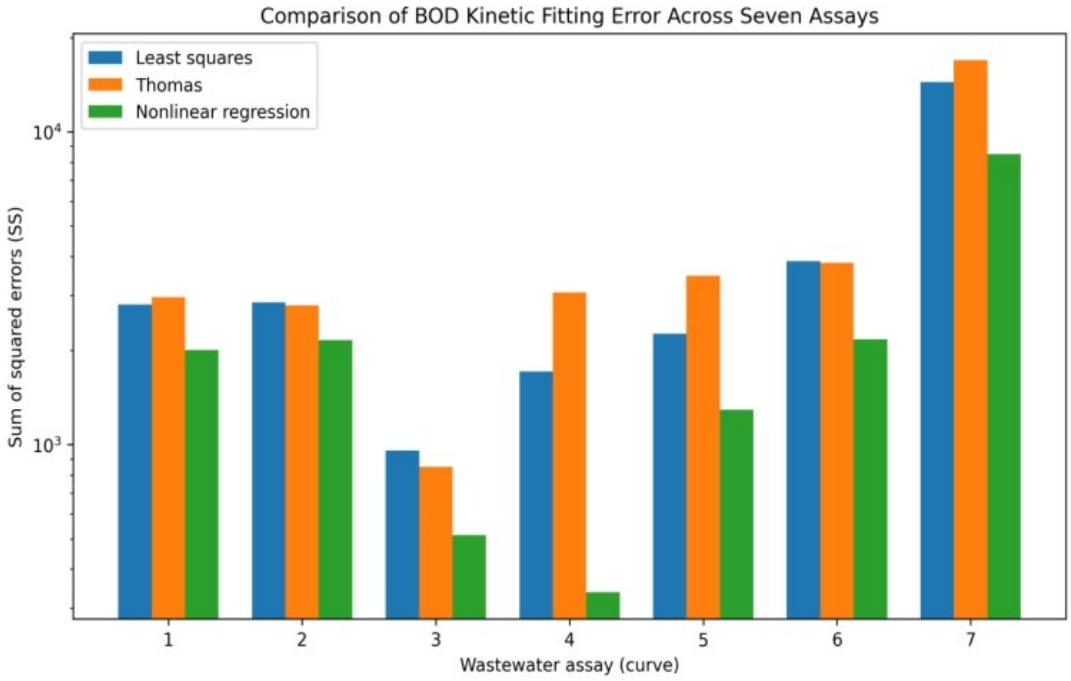


Figure 5: Squared Errors Comparison

Nonlinear regression produced the **lowest SS in all seven assays**. The mean SS was 2,423.03 for nonlinear regression, compared with 4,131.14 for least squares and 4,847.23 for the Thomas method. Relative to least squares, nonlinear regression reduced the mean SS by approximately:
4131.144131.14−2423.03×100 = 41.35%

Compared with the Thomas method:
4847.234847.23−2423.03×100 = 50.01%
Thus, nonlinear regression reduced the average squared fitting error by approximately **41.35% relative to least squares and 50.01% relative to Thomas**.

Table 7. Coefficient of determination (CD)

Curve	Least Squares	Thomas	Nonlinear Regression
1	0.99096	0.99047	**0.99355**
2	0.99407	0.99419	**0.99551**
3	0.99817	0.99838	**0.99902**
4	0.99642	0.99359	**0.99930**
5	0.99621	0.99415	**0.99783**
6	0.99402	0.99409	**0.99664**
7	0.99578	0.99556	**0.99770**
Mean	**0.99509**	**0.99435**	**0.99708**

Source: Cutrera et al. (1999), Table 2.

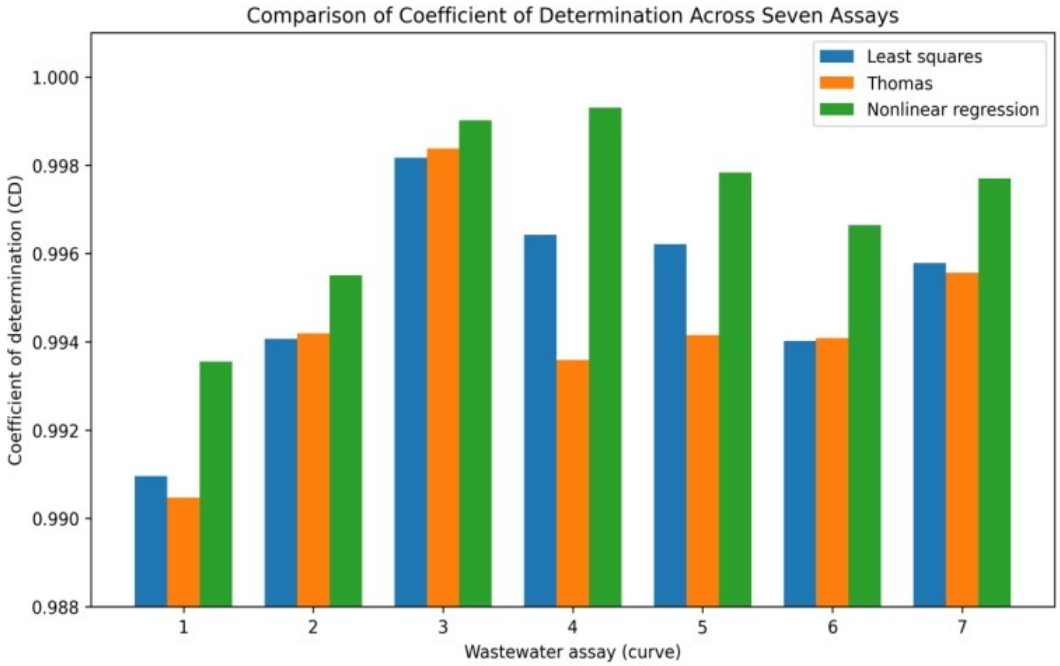


Figure 6: Coefficient of Determination Comparison
Nonlinear regression again produced the highest CD for every assay.

The mean CD increased from 0.99509 for least squares and 0.99435 for Thomas to 0.99708 for nonlinear regression.
The consistently higher CD indicates that nonlinear regression provided a closer representation of the variation in the observed BOD data.

Table 8. Model selection criterion (MSC)

Curve	Least Squares	Thomas	Nonlinear Regression
1	4.04	3.39	4.38
2	4.46	4.48	4.74
3	5.64	5.76	6.26
4	4.96	4.38	6.59
5	4.91	4.48	5.47
6	4.45	4.46	5.03
7	4.80	4.75	5.41
Mean	4.751	4.529	5.411

Source: Cutrera et al. (1999), Table 2.

Nonlinear regression achieved the highest MSC in all seven assays.

Its mean MSC of 5.411 was higher than both least squares (4.751) and Thomas (4.529).
The result is important because MSC provides an additional model-selection perspective beyond SS and CD. The original authors noted that MSC is not dependent on the numerical scale of the measurements and incorporates a penalty associated with model complexity.

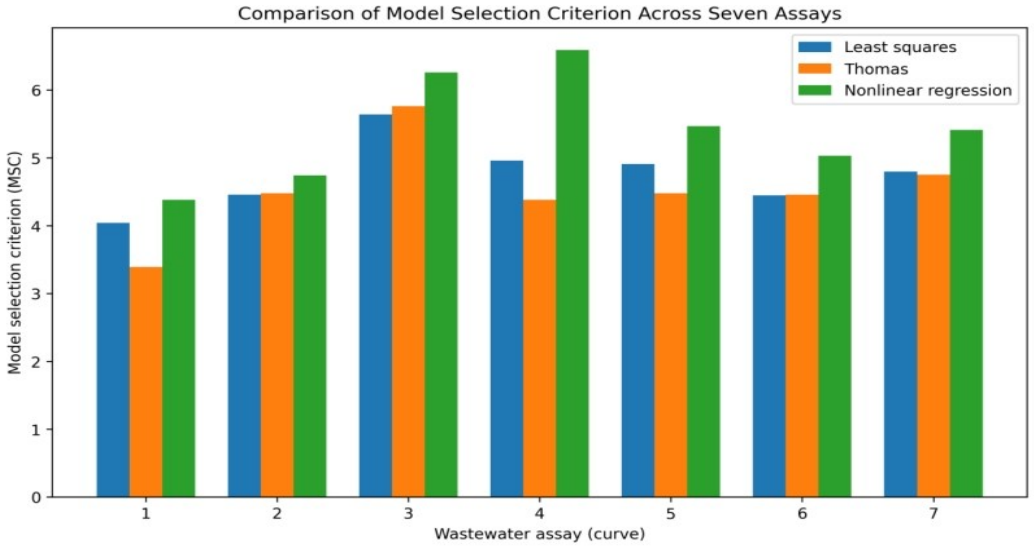


Figure 7: Model Selection Criterion Comparison

Table 9. Overall performance of the three BOD estimation methods

Performance indicator	Least Squares	Thomas	Nonlinear Regression
Mean SS	4,131.14	4,847.23	2,423.03
Mean CD	0.99509	0.99435	0.99708
Mean MSC	4.751	4.529	5.411
Mean k (day^{-1})	0.4811	0.4214	0.4722
Mean L0 (mg/L)	956.86	1,003.72	942.91
Best SS ranking	2nd	3rd	1st
Best CD ranking	2nd	3rd	1st
Best MSC ranking	2nd	3rd	1st

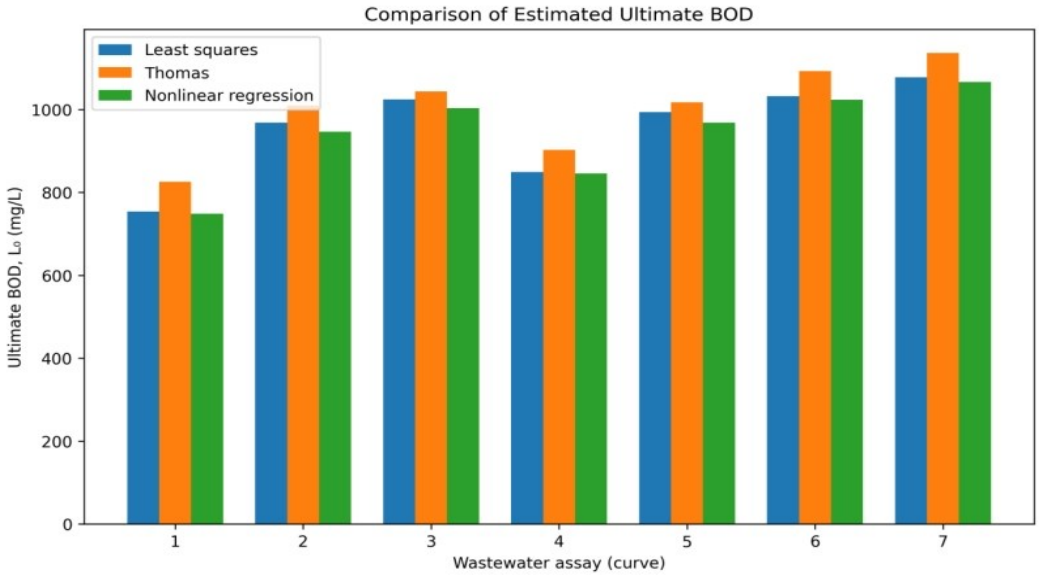


Figure 8: Comparison of Estimated Ultimate BOD

The results provide a consistent ranking:
Nonlinear Regression > Least Squares > Thomas
for the overall multi-criteria performance.
Importantly, this ranking is based on the direction of improvement appropriate to each criterion: lower SS and higher CD and MSC.

Discussion

The central result is that nonlinear regression was not merely better on average; it was the preferred method on every one of the seven curves under all three published goodness-of-fit criteria. This consistency is important because a method can have a favourable mean performance while failing substantially on individual cases. Here, the method ranking was uniform across the seven assays, which explains the large Kendall's W and the significant Friedman result.

The lower SS obtained by nonlinear regression indicates closer agreement between calculated and observed BOD values. The corresponding higher CD indicates that a larger fraction of the observed variation was represented by the fitted curve. MSC provides an additional model-selection perspective because it balances fit against parameterization. The simultaneous superiority of nonlinear regression on all three metrics strengthens the conclusion compared with relying only on R^2 or CD.

The findings are also consistent with the broader contemporary modelling literature, which continues to emphasize formal comparison of alternative predictive models for BOD and other wastewater-quality indicators. Recent studies have compared regression, machine-learning and ensemble approaches using multiple performance metrics, while current work on BOD prediction demonstrates the growing importance of rapid, reliable alternatives to conventional laboratory measurements. However, the present study addresses a different question: it compares three classical kinetic parameter-estimation procedures applied to the same BOD curves rather than predicting BOD from independent wastewater-quality variables.

The parameter estimates also show why numerical agreement in k and L_o should not be used alone to select an estimation method. Nonlinear and least-squares methods produced relatively close mean k and L_o values, whereas the Thomas method produced lower k and higher L_o. These differences arise from the transformations and fitting assumptions embedded in the respective procedures. In engineering applications, the preferred method should therefore be selected based on both parameter interpretability and goodness-of-fit performance.

Conclusions

A reproducible secondary-data analysis of seven published wastewater BOD assays was conducted to reassess least-squares, Thomas and nonlinear regression methods for estimating BOD kinetic parameters. Nonlinear regression consistently provided the lowest fitting error, the highest coefficient of determination and the highest model selection criterion on every curve. Mean SS was approximately 41% lower than least squares and 50% lower than Thomas; the mean CD and MSC were also higher. Formal paired analysis supported these differences, with Friedman $\chi^2(2) = 10.571$, $p = 0.005$ and Kendall's $W = 0.755$. Exact Wilcoxon tests indicated significant nonlinear-regression advantages over both alternatives after Bonferroni adjustment. The evidence therefore supports nonlinear regression as the most robust of the three evaluated procedures for the seven published assays. Future work should obtain raw BOD time-series data and compare classical kinetic models with modern statistical and machine-learning approaches using independent validation, uncertainty analysis and cross-wastewater generalization. It showed that nonlinear regression gave better values for SS, CD and MSC for every curve. The study addresses this gap by applying paired non-parametric inference and an effect-size

measure to the summary metrics. The study also quantifies the average percentage improvement of nonlinear regression over the two alternative methods.

References

Abd, S. A., & Al-Sulttani, A. O. (2024). Predicting biochemical oxygen demand at the inlet of Al- Rustumiya wastewater treatment plant using different mathematical techniques. Journal of Engineering. https://doi.org/10.31026/j.eng.2024.02.02

Adibimanesh, B., Polesek-Karczewska, S., Bagherzadeh, F., Szczuko, P., & Shafighfard, T. (2023) Energy consumption optimization in wastewater treatment plants: machine learning for monitoring incineration of sewage sludge. Sus. Ene. Techn. Assess., 56 , Article 103040

Adrian, D. D., & Sanders, T. G. (1992). Oxygen sag equation for half order BOD kinetics. Journal of Environmental Systems, 22(4), 341–345.

Aghdam, E., Mohandes, S. R., Manu, P., Cheung, C., Yunusa-Kaltungo, A., & Zayed, T. (2023). Predicting quality parameters of wastewater treatment plants using artificial intelligence techniques. Journal of Cleaner Production, 405, 137019. https://doi.org/10.1016/j.jclepro.2023.137019

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic *Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Akaike, H. (1976). An information criterion. Mathematical Science, 14(1), 5–9.

APHA, AWWA, & WEF. (2017). American Public Health Association (APHA), American Water Works Association (AWWA), & Water Environment Federation (WEF). (2017). Standard methods for the examination of water and wastewater (23rd ed.). Washington, DC: APHA. The Standard Methods series includes Method 5210 for biochemical oxygen demand, including the 5-day BOD and ultimate BOD procedures.

Brose, D. A. (2023). Assessment of statistical methods for converting biochemical oxygen demand and carbonaceous biochemical oxygen demand to total organic carbon in wastewater. Water and Environment Journal. https://doi.org/10.1111/wej.12834

Ching, P. M. L., Zou, X., Wu, D., So, R. H. Y., & Chen, G. H. (2022). Development of a wide-range soft sensor for predicting wastewater BOD_5 using an eXtreme gradient boosting (XGBoost) machine. Environmental Pollution.

Cutrera, G., Manfredi, L., del Valle, C. E., & González, J. F. (1999). On the determination of the kinetic parameters for the BOD test. Water SA, 25(3), 377–380.

Inbar, O., Shahar, M., & Avisar, D. (2024). Predictive modeling of BOD throughout wastewater treatment: a generalizable machine learning approach for improved effluent quality. Environmental Science: Water Research & Technology. https://doi.org/10.1039/D4EW00111G

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. SIAM Journal on Applied Mathematics, 11, 431–441.

Mathur, R., Sharma, M. K., Loganathan, K., et al. (2024). Modeling of two-stage anaerobic onsite wastewater sanitation system to predict effluent soluble chemical oxygen demand through machine learning. Scientific Reports, 14, 1835. https://doi.org/10.1038/s41598-023-50805-x

Metcalf & Eddy, Inc. (1977). Wastewater Engineering: Treatment, Disposal and Reuse. McGraw-Hill.

Moore, E. W., Thomas, H. A., Jr., & Snow, W. B. (1950). Simplified method for analysis of BOD data. Sewage and Industrial Wastes, 22, 1343–1355.

Moretti, A., Ivan, H. L., & Skvaril, J. (2024). A review of the state-of-the-art wastewater quality characterization and measurement technologies. Journal of Water Process Engineering, 60, 105061. https://doi.org/10.1016/j.jwpe.2024.105061

Oke, I. A., Lukman, S., Amoko, J. S., & Fehintola, E. O. (2017). An evaluation of solutions to moment method of biochemical oxygen demand kinetics. Nigerian Journal of Technology.

Qambar, A. S., & Al Khalidy, M. M. (2022). Prediction of municipal wastewater biochemical oxygen demand using machine learning techniques: A sustainable approach. Process Safety and Environmental Protection, 168, 833–845. https://doi.org/10.1016/j.psep.2022.10.033

Siwiec, T., Kiedryńska, L., Abramowicz, K., & Rewicka, A. (2012). Analysis of chosen models describing the changes in BOD_5 in sewages. Environment Protection Engineering, 38(2). https://doi.org/10.5277/epe120206

Thomas, H. A., Jr. (1950). Graphical determination of BOD curve constants. Water & Sewage Works, 97, 123–128.

Appendix A.
Dataset used for the analysis

Curve	k_LS	k_Thomas	k_NL	L0_LS	L0_Thomas	L0_NL	SS_LS	SS_Thomas	SS_NL	CD_LS	CD_Thomas	CD_NL	MSC_LS	MSC_Thomas	MSC_NL
Curve 1	0.41700	0.32990	0.40161	753.61100	825.50000	748.04000	2808.80000	2962.50000	2005.20000	0.99096	0.99047	0.99355	4.04000	3.39000	4.38000

Curve 2	0.38600	0.34500	0.38959	967.88100	1009.22000	946.19000	2851.80000	2792.20000	2160.10000	0.99407	0.99419	0.99551	4.46000	4.48000	4.74000
Curve 3	0.37600	0.35790	0.38089	1023.71200	1042.98000	1003.40000	958.40000	850.30000	513.50000	0.99817	0.99838	0.99902	5.64000	5.76000	6.26000
Curve 4	0.66100	0.55560	0.62346	849.37500	902.53000	845.23000	1714.90000	3069.50000	337.20000	0.99642	0.99359	0.99930	4.96000	4.38000	6.59000
Curve 5	0.54500	0.52150	0.55783	993.81900	1017.06000	968.23000	2260.20000	3472.50000	1293.60000	0.99621	0.99415	0.99783	4.91000	4.48000	5.47000
Curve 6	0.51400	0.43320	0.49477	1032.31500	1092.53000	1023.00000	3864.90000	3818.30000	2171.00000	0.99402	0.99409	0.99664	4.45000	4.46000	5.03000
Curve 7	0.46900	0.40650	0.45732	1077.28700	1136.21000	1066.30000	14459.00000	16965.30000	8480.60000	0.99578	0.99556	0.99770	4.80000	4.75000	5.41000

Note: values are transcribed from the published Tables 1 and 2; no raw time-series observations were reconstructed.