REVSTAT – Statistical Journal
Volume 14, Number 2, April 2016, 101–118

ESTIMATION ASPECTS OF THE MICHAELIS–MENTEN MODEL

Authors: Thomas L. Toulias
– Boulevard du Souverain 292, Auderghem 1160, Brussels, Belgium
th.toulias@gmail.com

Christos P. Kitsos
– Department of Informatics, Technological Educational Institute of Athens, Greece
xkitsos@teiath.gr

Received: September 2015 Revised: February 2016 Accepted: February 2016

Abstract:

- This paper studies the Michaelis–Menten model (MM), which plays an important role in pharmacokinetics, from a theoretical as well as a computational point of view. An analytical method for the nonlinear least squares estimation of the MM is introduced. It is proved that the MM model has not a unique parameter estimation (through the nonlinear least squares), and there is not a unique optimal experimental design and might not have a unique D-optimal design. An iterative process, based on the Sequential approach, is also introduced and tested on various data sets for the MM model. A different approach is also discussed which provides an initial estimate that increases the convergence rate of the Fully Sequential approach. Several examples demonstrate the provided methods.

Key-Words:

- *Michaelis–Menten model; optimal design; nonlinear least squares; fully sequential method.*

AMS Subject Classification:

- 62K05, 93E24, 62H12.

1. INTRODUCTION

A general theory for enzyme kinetics was firstly developed by Michaelis and Menten [17] in their pioneering work, where the metabolism of an agent is described by a reaction rate. The basic toxicokinetic model of metabolism is a Michaelis–Menten (MM) model. Briefly, when an enzyme E is combined reversibly with a substrate S to form an enzyme-substrate complex ES, which can be dissociate or proceed to the product P, the following enzyme-substrate reaction scheme

$$E + S \underset{k_1}{\overset{k_2}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P \,, \tag{1.1}$$

is assumed, with k_1, k_2 and k_3 being the associated rate constants. We let $K_{\rm M} := (k_2 + k_3)/k_1$, known as the MM constant, and $V_{\max} := k_3 C_{\rm T}$, where $C_{\rm T}$ is the total enzyme concentration. Then, a plot of the initial velocity of reaction V against the concentration of substrate $C_{\rm S}$, will provide the MM rectangular hyperbola of the form

$$V \,=\, V(C_{\rm S}) \,=\, V(C_{\rm S};\boldsymbol{\theta}) \,:=\, \frac{V_{\max}\, C_{\rm S}}{K_{\rm M} + C_{\rm S}} \,, \tag{1.2}$$

where the parameters' vector $\boldsymbol{\theta} := (V_{\max}, K_{\rm M}) \in \Theta \subseteq \mathbb{R}^2$, and Θ being the parameter space which is assumed compact when sequential approaches are applied. It is understood that the appropriate estimate of $\boldsymbol{\theta}$ is very essential to eliminate the Risk on the enzyme-substrate reaction scheme as in (1.1). The hidden Risk is strongly related to the appropriate real estimate, as it is proved bellow, a number of estimates might exist, even not real. Therefore the estimate $\boldsymbol{\theta}$ clarifies the Risk Analysis for the toxicokinetic model of metabolism under investigation, so essential in pharmacokinetics.

In practice, we have n readings for the reaction's initial velocity $V_i := V(C_{{\rm S},i};\boldsymbol{\theta})$ corresponding to n substrate concentration values $C_{{\rm S},i}$, $i = 1, 2, ..., n$. That is, only the stochastic model of the form $y_i := V_i + e_i$, $i = 1, 2, ..., n$ is obtained, as the readings V_i are associated with noise; see Kitsos [14, 13]. In principle, $C_{\rm S} > 0$ and hence $K_{\rm M} > 0$, while usually the reaction velocity $V(C_{\rm S}) > 0$.

Different linear transformations have been suggested, [4, 1], to estimate (with a linear regression fit) the involved parameters, $V_{\max}$ and $K_{\rm M}$, as the input variable $C_{\rm S}$ and the response V are curvilinearly related. The usual linear transformations are: the Eadie–Hofstee (EH), the Hanes–Wolf (HW), the Lineweaver–Burk or "double reciprocal" (LB), and the "inverse" Eadie–Hofstee (iEH) linearizations, which are formulated by

$$V \,=\, V_{\max} - K_{\rm M}\frac{V}{C_{\rm S}} \,, \qquad \text{(EH)} \tag{1.3a}$$

$$\frac{C_{\rm S}}{V} \,=\, \frac{K_{\rm M}}{V_{\max}} + \frac{1}{V_{\max}}\, C_{\rm S} \,, \qquad \text{(HW)} \tag{1.3b}$$

$$\frac{1}{V} \;=\; \frac{1}{V_{\max}} + \frac{K_{\mathrm{M}}}{V_{\max}}\,\frac{1}{C_{\mathrm{S}}}\,, \qquad \text{(LB)} \tag{1.3c}$$

$$C_{\mathrm{S}} \;=\; -K_{\mathrm{M}} + V_{\max}\,\frac{C_{\mathrm{S}}}{V}\,, \qquad \text{(iEH)} \tag{1.3d}$$

respectively, with the double reciprocal being the most popular. The Hanes–Wolf linearization has been shown in a very early work by Dowd and Riggs [4], as the most efficient. Endvenyi and Chan [6] and Currie [3] discussed the heteroscedasticity in the MM model. Endvenyi and Chan [6] assumed that the error variance is proportional to the mean. To avoid heteroscedasticity problems in the "linearized" models (1.3a)–(1.3d) we should suggest to solve the model's Normal Equations and get the least square estimators when n readings of V and C_{S} are given, i.e. to solve

$$\sum_{i=1}^{n}\left(V_i - \frac{V_{\max}C_{\mathrm{S},i}}{K_{\mathrm{M}}+C_{\mathrm{S},i}}\right)\nabla_{\boldsymbol{\theta}}V(C_{\mathrm{S},i};\boldsymbol{\theta}) \;=\; 0\,, \tag{1.4}$$

with V as in (1.2), while

$$\nabla_{\boldsymbol{\theta}}V \;=\; \left(\frac{\partial V}{\partial V_{\max}},\,\frac{\partial V}{\partial K_{\mathrm{M}}}\right)^{\mathrm{T}} = \left(\frac{C_{\mathrm{S}}}{K_{\mathrm{M}}+C_{\mathrm{S}}},\,-\frac{V_{\max}K_{\mathrm{M}}C_{\mathrm{S}}}{(K_{\mathrm{M}}+C_{\mathrm{S}})^2}\right)^{\mathrm{T}}. \tag{1.5}$$

Hence,

$$\frac{\partial S}{\partial V_{\max}} \;=\; \sum_{i=1}^{n}\left(V_i - \frac{V_{\max}C_{\mathrm{S},i}}{K_{\mathrm{M}}+C_{\mathrm{S},i}}\right)\frac{C_{\mathrm{S},i}}{K_{\mathrm{M}}+C_{\mathrm{S},i}} \;=\; 0\,, \tag{1.6}$$

$$\frac{\partial S}{\partial K_{\mathrm{M}}} \;=\; \sum_{i=1}^{n}\left(V_i - \frac{V_{\max}C_{\mathrm{S},i}}{K_{\mathrm{M}}+C_{\mathrm{S},i}}\right)\frac{V_{\max}C_{\mathrm{S},i}}{(K_{\mathrm{M}}+C_{\mathrm{S},i})^2} \;=\; 0\,. \tag{1.7}$$

Then, the two Normal Equations can be easily obtained and solved numerically so that the estimates $\hat{\boldsymbol{\theta}} = (\hat{V}_{\max}, \hat{K}_{\mathrm{M}})$ are obtained. For a single observation the Fisher's information matrix is $(\nabla V)^{\mathrm{T}}(\nabla V)$. Therefore, the average-per-observation information matrix $\mathbf{M}(\boldsymbol{\theta},\xi)$ is evaluated as

$$\sigma^{-2}\,n\,\mathbf{M}(\boldsymbol{\theta},\xi) \;=\; \sum_{i=1}^{n}\begin{pmatrix} C_{\mathrm{S},i}^2\,\tau_i^2 & -V_{\max}C_{\mathrm{S},i}^2\,\tau_i^3 \\ -V_{\max}C_{\mathrm{S},i}^2\,\tau_i^3 & V_{\max}C_{\mathrm{S},i}^2\,\tau_i^4 \end{pmatrix}, \tag{1.8}$$

with $\tau_i = 1/(K_{\mathrm{M}}+C_{\mathrm{S},i})$, $i = 1, 2, ..., n$. Thus, the 2×2 variance-covariance matrix is approximately equal to

$$\mathbf{C} \;=\; \mathbf{C}(\hat{\boldsymbol{\theta}},\xi) \;=\; \big(n\mathbf{M}(\hat{\boldsymbol{\theta}},\xi)\big)^{-1}. \tag{1.9}$$

Hence, we can derive asymptotic approximate confidence intervals for the involved parameters, and we can work for optimally criteria, with the D-optimal design being the most applicable; see [13] for details. Notice that, due to the fact that the MM model is partially nonlinear,[9], the D-optimal design depends only on the K_{M} parameter; see [14] for details.

2. NONLINEAR LEAST SQUARES FOR THE MM MODEL

In this Section an analytical method for the Nonlinear Least Square (NLLS) estimation of the MM model is introduced and discussed. In particular, the following Theorem provides a compact analytic methodology for the "actual" NLLS estimation of the MM model's parameters.

Recall that the Sum of Squares of Errors *sse* is given by $sse = sse(\boldsymbol{\theta}) := \sum_{i=1}^{n}[V_i - V(C_{\mathrm{S},i};\boldsymbol{\theta})]^2$, $n > 2$, where $\boldsymbol{\theta}$ is a vector of the MM model parameters $(V_{\max}, K_{\mathrm{M}})$, while $(C_{\mathrm{S},i}), (V_i) \in \mathbb{R}^n$ are the data vectors for the substrate concentration and the reaction's velocity respectively. The estimated parameters $\hat{\theta}_1 = \hat{V}_{\max}$ and $\hat{\theta}_2 = \hat{K}_{\mathrm{M}}$, from the estimated vector $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2)$, are obtained when $sse(\hat{\boldsymbol{\theta}}) = \min\{sse(\boldsymbol{\theta})\}_{\boldsymbol{\theta}\in\Theta}$, $\Theta \in \mathbb{R}^2$ compact, i.e. when $\nabla S = 0$, or equivalently, when the normal equations (1.4) are satisfied. Recall also that the mean absolute relative error $mare = mare(\hat{\boldsymbol{\theta}}) := \mathrm{E}\big(\big|\{V_i - V(C_{\mathrm{S},i};\hat{\boldsymbol{\theta}})\}/V_i\big|\big)$, is also evaluated, see Example 2.1 below, while, due to [10], there is always a solution of the normal equations.

The following Theorem as it is stated and proved, provides evidence that the MM model has not unique Least Square Estimates. As we already mentioned in the introduction this creates a further investigation for the Risk Analysis under study, as the appropriate, among a number of estimates, has to be chosen. We discuss the proposed strategy in the sequence of this paper.

Theorem 2.1. *The NLLS estimators $\hat{V}_{\max}$ and $\hat{K}_{\mathrm{M}}$ of the MM model are the ones among the $K \le 4n-5$ possible estimates' vectors $\hat{\boldsymbol{\theta}}_k = (\hat{V}_{\max;k}, \hat{K}_{\mathrm{M};k})$, $k = 1, 2, ..., K$, with $sse(\hat{V}_{\max}, \hat{K}_{\mathrm{M}}) := \min\{sse(\hat{V}_{\max;k}, K_{\mathrm{M};k})\}_{k=1,2,...,K}$, where*

$$\hat{V}_{\max;k} = \left(\sum_{i=1}^{n} \frac{V_i C_{\mathrm{S},i}}{\hat{K}_{\mathrm{M};k} + C_{\mathrm{S},i}}\right)\left[\sum_{i=1}^{n}\left(\frac{C_{\mathrm{S},i}}{\hat{K}_{\mathrm{M};k} + C_{\mathrm{S},i}}\right)^2\right]^{-1}, \quad k = 1, 2, ..., K\,, \tag{2.1}$$

while the estimated $\hat{K}_{\mathrm{M};k}$ values are the $K \le 4n-5$ real roots of the following $(4n-5)$-degree polynomial of ϑ_2:

$$P(\vartheta_2) := \sum_{\substack{i,j=1\\(i\neq j)}}^{n} V_i C_{\mathrm{S},i} C_{\mathrm{S},j} (C_{\mathrm{S},j} - C_{\mathrm{S},i})(\vartheta_2 + C_{\mathrm{S},i})^2 (\vartheta_2 + C_{\mathrm{S},j}) \prod_{(i,j\neq)\,k=1}^{n} (\vartheta_2 + C_{\mathrm{S},k})^4\,. \tag{2.2}$$

Proof: Solving the first normal equation (1.6) with respect to $V_{\max}$ and substituting to the second one (1.6) we obtain

$$\left(\sum_{i=1}^{n} \frac{V_i x_i}{(K_{\mathrm{M}} + x_i)^2}\right)\sum_{i=1}^{n}\left(\frac{x_i}{K_{\mathrm{M}} + x_i}\right)^2 = \left(\sum_{i=1}^{n} \frac{V_i x_i}{K_{\mathrm{M}} + x_i}\right)\left(\sum_{i=1}^{n} \frac{x_i^2}{(K_{\mathrm{M}} + x_i)^3}\right),$$

where $x_i := C_{\mathrm{S},i}$, $i = 1, 2, ..., n$. Thus,

$$\left[\sum_{i=1}^{n} u_i x_i Q_i(2)\right]\left[\sum_{j=1}^{n} x_j^2 Q_j(2)\right] = \left[\sum_{i=1}^{n} u_i x_i Q_i(1)\right]\left[\sum_{j=1}^{n} x_j^2 Q_j(3)\right],$$

where $Q_i(d) := \prod_{(i\neq)\,m=1}^{n}(K_{\mathrm{M}} + C_{\mathrm{S},m})^d$, $i = 1, 2, ..., n$, $d = 1, 2, 3$. With some algebra, the above relation can be written as

$$\sum_{\substack{i,j=1\\(i\neq j)}}^{n} u_i x_i x_j^2 \left\{ s_i^2 s_j^2 \left(\prod_{(i,j\neq)\,m=1}^{n} s_m^4\right) - s_i^3 s_j \left(\prod_{(i,j\neq)\,m=1}^{n} s_m^4\right)\right\} = 0\,, \tag{2.3}$$

where $s_m := K_{\mathrm{M}} + C_{\mathrm{S},m}$, $m = 1, 2, ..., n$. Finally, the solution with respect to K_{M} of the above equation corresponds to the roots of the polynomial P of $\vartheta_2 := K_{\mathrm{M}}$, as in (2.2), as then the requested K_{M} $(= \vartheta_2)$ values would satisfy the normal equations (1.6) and (1.7). For each of the $K \leq 4n - 5$ real-valued root of (2.2), i.e. for each possible estimate $\hat{K}_{\mathrm{M};k}$, the corresponding $\hat{V}_{\max;k}$, $k = 1, 2, ..., K$, estimate is then obtained through (2.1), which is the solution of the (1.6) with respect to $V_{\max}$.

The solutions of the normal equations (1.6) and (1.7), through the roots of the polynomial (2.2), provide $4n-5$ possible estimates of $\boldsymbol{\theta}$. As this number is odd there is always at least one real root of (2.2). Therefore, at least one real critical point of the least square objective function S exists, which may yield at least one estimate $\hat{\boldsymbol{\theta}} = (\hat{\vartheta}_1, \hat{\vartheta}_2) := (\hat{V}_{\max}, \hat{K}_{\mathrm{M}})$ for the MM model. For a working example see [18]. However, when all data points are collinear, the nonlinear least squares estimate cannot exist, see [11] or [8]. Therefore, various different NLLS estimates may exist (among the K real-valued $\hat{\boldsymbol{\theta}}_k = (\hat{V}_{\max;k}, \hat{K}_{\mathrm{M};k})$, $k = 1, 2, ..., K$) which (locally) minimizes the sum of squares $S = sse(\hat{\boldsymbol{\theta}})$. The problem the experimenter has then to face is which of the real roots ϑ $(= \hat{K}_{\mathrm{M}})$ of $P = P(\vartheta)$, as in (2.2), can be chosen as the MM model's NLLS estimate $\hat{K}_{\mathrm{M}}$. Among the K real-valued (of the total $4n - 5$) candidate estimates, the experimenter can always choose the one which provides the minimum sum of squares, i.e.

$$sse(\hat{\boldsymbol{\theta}}_{\mathrm{NLLS}}) := sse(\hat{V}_{\max}, \hat{K}_{\mathrm{M}}) = \min\left\{sse(\hat{\boldsymbol{\theta}}_k)\right\}_{k=1,2,...,K}, \tag{2.4}$$

as the NLLS estimates' vector $\hat{\boldsymbol{\theta}}_{\mathrm{NLLS}}$ would then be a global minimum for the MM model's sum of squares. □

As the degree $(4n-5)$ of the polynomial $P(\vartheta_2)$ in Theorem 2.1 is odd, there is always at least one (real-valued) estimates' vector $\hat{\boldsymbol{\theta}}$ for the MM model. Thus, as more than one estimates's vectors can exist, then more than one corresponding average-per-observation information matrices are possible. Therefore, the design might not be considered as unique, as Biebler in [2] has been noticed. See also [16]. The criterion we suggest (the minimum *sse*) actually offers the design with the minimum variance, i.e. it corresponds to the D-optimal one.

The following Example provides a comparative presentation of the linear and the "actual" nonlinear estimation for the MM model.

Example 2.1. The treated case of the Puromicin data set was adopted, as in Bates and Watts [1, Table A1.3], for the comparison of the Linear Least Squares (LLS) estimations, through the linearizations as in (1.3a)–(1.3d), and the analytic NLLS estimation of Theorem 2.1. Table 1 provides the evaluated LLS estimates $\hat{V}_{\max}$ and $\hat{K}_{\mathrm{M}}$ (obtained by linear regression), which correspond to the Eadie–Hofstee (EH), Hanes–Wolf (HW), Lineweaver–Burk (LB), and the "inverse" Eadie–Hofstee (iEH) linearization methods, together with the NLLS estimates that provide the minimum sum of squares (among all the possible pairs of NLLS estimates obtained through Theorem 2.1). Their R^2 coefficient as well as their corresponding *sse* and *mare* (%) errors are also presented. All the involved calculations in Table 1 as well as the corresponding Figure 1 were done by using MATLAB as a programming tool.

Table 1: Comparison between the LLS and the analytic NLLS estimation.

Est. Method	$\hat{V}_{\max}$	$\hat{K}_{\mathrm{M}}$	R^2	*sse*	*mare* (%)
LLS (EH)	193.867711	0.043524	0.9282	184.69	10.74
LLS (HW)	216.216899	0.067915	0.9603	102.05	6.94
LLS (LB)	195.802709	0.048407	0.9378	160.05	9.19
LLS (iEH)	215.773203	0.067068	0.9606	101.45	6.99
NLLS	212.683743	0.064121	0.9613	99.62	6.99

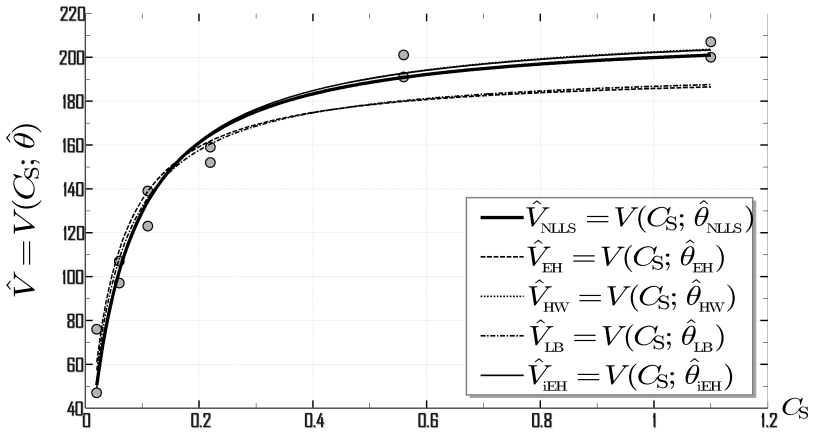

Figure 1: Visual comparison between the predicted NLLS model $\hat{V}_{\mathrm{NLLS}}$ against the four predicted LLS models.

It is clear that the NLLS estimation provides a "better" estimate than the LLS ones, in terms of the corresponding R^2 coefficients, the sum of squared errors

sse and the mean absolute relative errors *mare*. Notice that, the estimation through the iEH linearization approximates better the analytic NLLS estimation, as adopts the least sum of squared error *sse*, and almost identical *mare* error among the other three LLS estimation.

Figure 1 provides a graphic comparison between the linear and the nonlinear least squares estimation for the MM model (using the Puromicin data set), by depicting the estimated NLLS model $\hat{V}_{\mathrm{NLLS}} := V(C_{\mathrm{S}};\hat{\boldsymbol{\theta}}_{\mathrm{NLLS}})$ against with the four LSS estimated models $\hat{V}_{\mathrm{EH}} := V(C_{\mathrm{S}};\hat{\boldsymbol{\theta}}_{\mathrm{EH}})$, $\hat{V}_{\mathrm{HW}} := V(C_{\mathrm{S}};\hat{\boldsymbol{\theta}}_{\mathrm{HW}})$, $\hat{V}_{\mathrm{LB}} := V(C_{\mathrm{S}};\hat{\boldsymbol{\theta}}_{\mathrm{LB}})$ and $\hat{V}_{\mathrm{iEH}} := V(C_{\mathrm{S}};\hat{\boldsymbol{\theta}}_{\mathrm{iEH}})$. The estimates' vectors $\hat{\boldsymbol{\theta}}_{\mathrm{NLLS}}$, $\hat{\boldsymbol{\theta}}_{\mathrm{EH}}$, $\hat{\boldsymbol{\theta}}_{\mathrm{HW}}$, $\hat{\boldsymbol{\theta}}_{\mathrm{LB}}$ and $\hat{\boldsymbol{\theta}}_{\mathrm{iEH}}$ are provided by the corresponding vectors $(\hat{V}_{\max},\hat{K}_{\mathrm{M}})$ of Table 1. Note that the iEH and the HW linearizations provide almost similar LLS models (the dotted $\hat{V}_{\mathrm{HW}}$ curve is very close to the thin solid $\hat{V}_{\mathrm{iEH}}$ curve), both very close to the 'actual' NLLS model.

3. SEQUENTIAL GAUSS–NEWTON ESTIMATION

The Fully Sequential (FS) method has been discussed by Ford *et al.* in [7] and Kitsos in [15, 12] for the nonlinear experimental design problem. Expanding the FS method, in this Section we introduce and investigate the general case of the Batch Sequential (BS) iterative scheme, for the MM model estimation.

For the estimation of the parameter $\boldsymbol{\theta}\in\Theta$, $\Theta\subseteq\mathbb{R}^q$ compact, of any model $\eta=\eta(x;\boldsymbol{\theta})$ in general, recall the known definition of sum of squares $sse = sse(\boldsymbol{\theta})$,

$$sse(\boldsymbol{\theta}) = sse(\boldsymbol{\theta};\mathbf{x}) := \big\|\boldsymbol{\eta}-\eta(\mathbf{x};\boldsymbol{\theta})\big\|^2 = \sum_{i=1}^{n}\big[\eta_i-\eta(x_i;\boldsymbol{\theta})\big]^2\,, \tag{3.1}$$

where $\mathbf{x}:=(x_i)\in\mathbb{R}^n$, $\boldsymbol{\eta}:=(\eta_i)\in\mathbb{R}^n$, and $\|\cdot\|$ denotes the usual $\mathscr{L}^2$-norm. The estimated parameters' vector $\hat{\boldsymbol{\theta}}=(\hat{\vartheta}_1,\hat{\vartheta}_2)$ is obtained when $sse(\hat{\boldsymbol{\theta}})=\min\{sse(\boldsymbol{\theta})\}_{\boldsymbol{\theta}\in\Theta}$, i.e. when $\nabla S=0$.

Recall also the iterative GN method, for the parameter estimation of the general nonlinear model described by $\eta=\eta(x;\boldsymbol{\theta})$; see also [5]. In the GN iterative procedure a series of estimates $\hat{\boldsymbol{\theta}}_N\in\Theta$ is produced where the next estimates' vector $\hat{\boldsymbol{\theta}}_{N+1}$ is derived from the previous one $\hat{\boldsymbol{\theta}}_N$. When the sequence converges to a vector, then this vector is a possible NLLS estimates' vector, say $\hat{\boldsymbol{\theta}}_{\mathrm{NLLS}}\in\Theta$, of the model η, i.e. $\nabla sse(\hat{\boldsymbol{\theta}}_{\mathrm{NLLS}})=0$. The GN iterative scheme is described in a compact form, by

$$\hat{\boldsymbol{\theta}}_{N+1} = \hat{\boldsymbol{\theta}}_N - \mathbf{H}_S^{-1}(\hat{\boldsymbol{\theta}}_N)\,\nabla sse(\hat{\boldsymbol{\theta}}_N)\,,\qquad N\in\mathbb{N}^*:=\mathbb{N}\backslash\{0\}\,, \tag{3.2}$$

for a given initial estimates' vector $\hat{\boldsymbol{\theta}}_0=(\hat{\vartheta}_1^0,\hat{\vartheta}_2^0)\in\Theta$, where $\mathbf{H}_S^{-1}(\hat{\boldsymbol{\theta}}_N)$ is the inverse of the Hessian matrix of the sum of squares function $sse(\hat{\boldsymbol{\theta}}_N)$ as in (3.1).

The k-Batch Sequential approach (k-BS), $k \in \mathbb{N}^*$, presented here, is based on the GN method, and it consists of N_k steps in total. On the N-th step of the k-BS method, a number of GN iterations are performed on the chunk of Nk observations (from the total n), which always starts form the first and ends to the Nk-th observation. In particular, the total number of iterations N_k is given by

$$N_k := \begin{cases} n/k\,, & \text{when } \ n/k \in \mathbb{N}^*, \ \text{ and } \ k \neq 1\,, \\ [n/k]+1\,, & \text{when } \ n/k \notin \mathbb{N}^*, \ \text{ and } \ k \neq 1\,, \\ n-1\,, & \text{when } \ k = 1\,, \end{cases} \tag{3.3}$$

where $[\cdot]$ denotes the integer part of a real number and n is the number of observations. Note that, for the application of the BS iterative process, the data pairs $\{(x_i, \eta_i)\}_{i=1,2,...,n}$ are "entered" sequentially into the BS process. Therefore, instead of the usual sum of squares S, as in (3.1), the k-BS approach utilizes a partial sum of squares (p.s.s.) S_N (used for the corresponding GN iterations on every step N of the k-BS method), which is the sum of squares calculated only for the specific chunk of Nk observations on the N-th step of the method. Thus, the S_N is defined as

$$S_N(\hat{\boldsymbol{\theta}}) := \sum_{i=1}^{kN} \Big[V_i - V(C_{\mathrm{S},i}; \hat{\boldsymbol{\theta}})\Big]^2, \qquad N = 1, 2, ..., N_k = n/k\,, \tag{3.4}$$

provided that $n/k \in \mathbb{N}$, with $k \neq 1$. For the case of $n/k \notin \mathbb{N}^*$ (and $k \neq 1$), the p.s.s. S_N is defined as in (3.4) for $N = 1, 2, ..., [n/k]$, while for the last step $N_k = [n/k] + 1$, the $S_{N=N_k}$ is defined by

$$S_N(\hat{\boldsymbol{\theta}}) := \sum_{i=1}^{n} \Big[V_i - V(C_{\mathrm{S},i}; \hat{\boldsymbol{\theta}})\Big]^2. \tag{3.5}$$

For the special case of the 1-BS method, the p.s.s. S_N is defined by

$$S_N(\hat{\boldsymbol{\theta}}) := \sum_{i=1}^{N+1} \Big[V_i - V(C_{\mathrm{S},i}; \hat{\boldsymbol{\theta}})\Big]^2, \qquad N = 1, 2, ..., N_1 = n-1\,. \tag{3.6}$$

Finally, the GN iterative procedures at each step N (of the total N_k steps) of the k-BS method, either converge or can be stopped (after a given maximum number of GN iterations) to some estimated parameters' vector, say $\hat{\boldsymbol{\theta}}_N$. This $\hat{\boldsymbol{\theta}}_N$ is then considered as the initial vector for the GN iterations of the next step $N+1$ of the k-BS method. Hence, an initial parameters' vector $\hat{\boldsymbol{\theta}}_0$ is then needed in order to begin the GN iterations of first step $N = 1$ of the k-BS method.

Notice that, for a set of $n = mk$, $m \in \mathbb{N}^*$, observations, the p.s.s. S_N (for the GN iterations on the N-th step of the k-BS scheme, $k \neq 1$) is calculated through the summation of successively $k, 2k, 3k, ..., mk = n$ terms. For a set of odd number of observations, say $n = mk + q$ with $\mathbb{N}^* \ni q < m$, the p.s.s. S_N

is calculated through the summation of successively $k, 2k, 3k, ..., mk, n$ terms. For the 1-BS case, the corresponding p.s.s. S_N summation is performed with successively $2, 3, ..., n$ terms. Notice also that the 2-BS scheme coincides with the FS scheme, as the p.s.s. S_N summation uses $2, 4, 6, ..., n$ terms.

The above description of the k-BS iterative method can be formulated into the following algorithm.

Algorithm 3.1. Consider an initial vector $\hat{\boldsymbol{\theta}}_0$ for the estimation of the model $\eta = \eta(x; \boldsymbol{\theta}) = \eta(x; \vartheta_1, \vartheta_2)$. On every step $N = 1, 2, ..., N_k$ of the k-BS method, a GN iterative process is applied, using the p.s.s. S_N for a given maximum number of iterations, say J. Then, a series of vectors is produced, say $\hat{\boldsymbol{\theta}}_{N,1}, \hat{\boldsymbol{\theta}}_{N,2}, ..., \hat{\boldsymbol{\theta}}_{N,J}$. The next estimate $\hat{\boldsymbol{\theta}}_{N+1}$ is then considered to be the last current estimate

$$\hat{\boldsymbol{\theta}}_{N+1} = \hat{\boldsymbol{\theta}}_{N,J}\,, \qquad N = 0, 1, 2, ..., N_k\,,$$

where the vectors $\hat{\boldsymbol{\theta}}_{N,j}$, $j = 1, 2, ..., J$, are described by the GN iterative scheme

$$\hat{\boldsymbol{\theta}}_{N,j+1} = \hat{\boldsymbol{\theta}}_{N,j} - \mathbf{H}^{-1}_{S_N}(\hat{\boldsymbol{\theta}}_{N,j})\nabla S_N(\hat{\boldsymbol{\theta}}_{N,j})\,, \quad j = 0,1, ..., J_N \leq J\,, \text{ with } \hat{\boldsymbol{\theta}}_{0,0} := \hat{\boldsymbol{\theta}}_0\,. \tag{3.7}$$

For every step N, the index $J_N \leq J$ is the one for which the GN process converges (when it does), i.e. when the convergence error, say e_N, of the estimate $\hat{\boldsymbol{\theta}}_{N,J_N} = (\hat{\vartheta}^1_{N,J_N}, \hat{\vartheta}^2_{N,J_N})$ is smaller or equal than a given threshold error e, i.e.

$$e_N := \max\Big\{\big|\hat{\vartheta}^1_{N,J_N} - \hat{\vartheta}^1_{N,J_N-1}\big|,\, \big|\hat{\vartheta}^2_{N,J_N} - \hat{\vartheta}^2_{N,J_N-1}\big|\Big\} \leq e\,, \quad N = 1, 2, ..., N_k\,. \tag{3.8}$$

Otherwise, when the convergence fails, the GN process stops at the J-th GN iteration (i.e. when $j = J$). For the next $N + 1$ step, as an initial vector $\hat{\boldsymbol{\theta}}_{N+1,0}$ for the new GN iteration, we consider the last estimate of the previous GN process, i.e. $\hat{\boldsymbol{\theta}}_{N+1,0} = \hat{\boldsymbol{\theta}}_{N,J_N}$, $N = 1, 2, ..., N_k$.

The following Example applies the FS iterative scheme, i.e. the 2-BS iterative scheme as in Algorithm 3.1.

Example 3.1. The Puromycin-treated data set, as in Example 2.1, consists of $n = 12$ observations where the $C_{\mathrm{S},i}$, $i = 1, 2, ..., 6$, readings are repeated. For this Example we consider the subset of the $n = 6$ non-replicated observations of the Puromycin data set. Let $e = 10^{-4}$ be the convergence error threshold of the 2-BS method, while as initial estimates' vector guess we adopt $\hat{\boldsymbol{\theta}}_0 = (100, 0)$. The first sub-Table of the Table 2 provides the last J_N-th GN convergent estimates' vector $\hat{\boldsymbol{\theta}}_{N,J_N}$, for each of the (total three) steps $N = 1, 2, 3$ of the 2-BS method. Recall that the total number of steps of the 2-BS algorithm for this data set is $N_{k=2} := n/k = 6/2 = 3$. Notice that $0 + 7 + 7 + 5 = 19$ GN iteration are needed, in total, to obtain $\hat{\boldsymbol{\theta}}_{\mathrm{NLLS}}$ with accuracy $< 10^{-5}$. Moreover, the GN processes, for every step N of the 2-BS scheme, do not have to converge at a

given error convergence e. For example, reducing the maximum number of the GN iterations to be, say $J=5$, we obtain also $\hat{\boldsymbol{\theta}}_{\mathrm{NLLS}}$ with the same accuracy $<10^{-5}$, but this time $0+5+5+5=15$ GN iteration are needed in total, see the middle sub-Table of Table 2.

Table 2: Convergence of various 2-BS processes, for the MM model estimation.

N	J_N	$\hat{V}_{\max}$	$\hat{K}_{\mathrm{M}}$	e_N	ε_N	R^2	*mare* (%)
• *Maximum # of GN iterations:* $J=10$							
0	0	100.	0.	—	2.5e+5	−1.6531	17.34
1	7	112.549618	0.009618	3.89e−8	0.	1.	0.
2	7	172.241997	0.034754	5.36e−10	7.64e−11	0.8739	10.63
3	5	**210.857219**	**0.062575**	7.03e−6	1.01e−8	0.9360	9.341
• *Maximum # of GN iterations:* $J=5$							
0	0	100.	0.	—	2.5e+5	−1.6531	17.34
1	5	112.548698	0.009618	0.143	2.18	1.	0.00043
2	5	172.241882	0.034754	0.0539	0.218	0.8739	10.63
3	5	**210.857219**	**0.062575**	7.03e−6	1.01e−8	0.9360	9.341
• $\hat{\boldsymbol{\theta}}_0=(V_{\max},K_{\mathrm{M}})$, $J=10$							
0	0	112.549618	0.009618	—	0.	1.	0.
1	1	112.549618	0.009618	0.	0.	1.	0.
2	7	172.241997	0.034754	5.36e−10	7.64e−11	0.8739	10.63
3	5	**210.857219**	**0.062575**	7.03e−6	1.01e−8	0.9360	9.341

The initial estimate guess the experimenter provides, plays a crucial role for the convergence of the general BS methodology. In order to address this issue, the 2-BS scheme can be applied adopting as initial parameters' vector $\hat{\boldsymbol{\theta}}_0$ the solution vector $(V_{\max},K_{\mathrm{M}})$ of the two MM model relations $V_i=V(C_{\mathrm{S},i};V_{\max},K_{\mathrm{M}})$, $i=1,2$. These relations can be solved analytically in the form of $(V_{\max},K_{\mathrm{M}})=(V_1+dV_1,\ dC_{\mathrm{S},2})$, $d:=C_{\mathrm{S},2}(V_2-V_1)/(V_1C_{\mathrm{S},2}-V_2C_{\mathrm{S},1})$. As a result, the first GN process (for the step $N=1$ of the 2-BS approach) will always converge at its first GN iteration $j=1=J_1$. This is due to the fact that the initial GN iteration process (where only the first two observations are involved) will surely converges to the only solution $(V_{\max},K_{\mathrm{M}})$, as above, which is now provided by the suggested initial vector $\hat{\boldsymbol{\theta}}_0$. This suggested $\hat{\boldsymbol{\theta}}_0$ it turns to be the convergent estimate $\hat{\boldsymbol{\theta}}_{1,J_1}$ of the first (and only) GN iteration for first step $N=1$, of the 2-BS process. Therefore, with the above proposed $\hat{\boldsymbol{\theta}}_0$ there is no need for guessing an initial vector, to feed the 2-BS process, that might not converge. This is true at least at the first step of the 2-BS approach. The last sub-Table of Table 2 provides the last J_N-th GN convergent estimates $\hat{\boldsymbol{\theta}}_{N,J_N}$ for every step $N=1,2,3$ of the 2-BS process, adopting as $\hat{\boldsymbol{\theta}}_0$ the proposed solution $(V_{\max},K_{\mathrm{M}})$ as discussed above. Table 2 also presents the convergence error e_N, the solution error distance $\varepsilon_N:=\|\nabla sse(\hat{\boldsymbol{\theta}}_N)\|$ as well as the R^2 coefficient for each estimated model $\hat{V}_N:=V(C_{\mathrm{S}};\hat{\boldsymbol{\theta}}_{N,J_N})$. The digits in bold represents the accurate digits of the NLLS estimates.

From our experience, due to the sequential nature of the BS methodology, the behaviour of the BS approach it might depend on the order in which the data entered into the sequential process. As the MM model is a strictly monotonous function, it is generally preferred that the $C_{\mathrm{S},i}$ readings of the data set $\{(C_{\mathrm{S},i}, V_i)\}_{i=1,2,\ldots,n}$ are maintain this strictly monotonic pattern as they entered sequentially into the BS algorithm. However, the problem might arises when the BS process is applied on a "replicated" data set, i.e. a data set showing multiple values (two or more) of V_i for each $C_{\mathrm{S},i}$ value. Such data set is the Puromycin-treated data set which consisted of two V_i's readings for every (out of six) $C_{\mathrm{S},i}$ value. In particular, the problem is occurred when we adopt the MM solution vector $(V_{\max}, K_{\mathrm{M}})$ to play the role of the initial estimates' vector $\hat{\boldsymbol{\theta}}_0$, as we suggested earlier. Unfortunately, the solution vector $(V_{\max}, K_{\mathrm{M}})$ of the two MM relations $V_i = V(C_{\mathrm{S},i}; V_{\max}, K_{\mathrm{M}})$, $i = 1, 2$, does not exist (as $C_{\mathrm{S},1} = C_{\mathrm{S},2}$ in this data set). To avoid this problem we can apply a "higher order" BS process, as the 4-BS.

The following Example demonstrates the above discussion.

Example 3.2. The 4-BS iterative process is applied for the Puromycin data set as well as for the Enzyme Velocity (EV) data set in [11, pg. 242], which are both having replicated (double) values of C_{S} readings. Let $e = 10^{-4}$ be the threshold for the convergence error. For an initial guess $\hat{\boldsymbol{\theta}}_0$ we obtain firstly all the solutions $(V_{\max}, K_{\mathrm{M}})$ between the two MM model relations $V_i = V(C_{\mathrm{S},i}; V_{\max}, K_{\mathrm{M}})$ and $V_j = V(C_{\mathrm{S},j}; V_{\max}, K_{\mathrm{M}})$, for all $i \neq j$, $i, j = 1, 2, ..., n$. Then we adopt as $\hat{\boldsymbol{\theta}}_0$ that specific solution vector $(V_{\max}, K_{\mathrm{M}})$ which provides the minimum sum of squares for the corresponding MM model, i.e.

$$sse(\hat{\boldsymbol{\theta}}_0) = \min\Big\{sse(\boldsymbol{\theta}):\ \boldsymbol{\theta} \text{ is the solution of } V_k = V(C_{\mathrm{S},k}; \boldsymbol{\theta}),\ k \in \{i,j\}\Big\}_{i,j\in\{1,2,3,4\}},$$

or equivalently

$$sse(\hat{\boldsymbol{\theta}}_0) = \min\Big\{sse(\vartheta_1, \vartheta_2):\ \vartheta_1 = V_i(1 + d_{ij}),\ \vartheta_2 = C_{\mathrm{S},i}\, d_{ij}\Big\}_{i,j\in\{1,2,3,4\}}, \tag{3.9}$$

where $d_{ij} := C_{\mathrm{S},j}(V_j - V_i)/(V_i C_{\mathrm{S},j} - V_j C_{\mathrm{S},i})$, $i, j = 1, 2, 3, 4$. As the initial estimate $\hat{\boldsymbol{\theta}}_0$ is used for the application of the 4-BS process, it should to be obtained by using the first 4 observations (recall that the GN process at the first step of the 4-BS method is performed using the first four observations), and thus $\hat{\boldsymbol{\theta}}_0$ should satisfy (3.9) where $n := 4$.

Table 3 presents the results of the 4-BS approach for the Puromycin and the EV data sets, where the presented estimates $\hat{\boldsymbol{\theta}}_{N,J_N} = (\hat{V}_{\max}, \hat{K}_{\mathrm{M}})$ are calculated with only $J = 5$ maximum number of GN iterations on every step $N = 1, 2, 3$ (of the 4-BS algorithm). Notice the remarkable accuracies, of less than $10^{-7} < e$ (achieved in total $0 + 5 + 5 + 5 = 15$ GN iterations) for the requested $\hat{\boldsymbol{\theta}}_{\mathrm{NLLS}}$, for

the Puromycin data set, and of less than $10^{-5} < e$ for the EV data set (achieved in total $0+8+8+5=21$ GN iterations). The digits in bold represents the accurate digits of the NLLS estimates.

Table 3: Convergence of the 4-BS processes, for the MM model estimation.

N	J_N	$\hat{V}_{\max}$	$\hat{K}_{\mathrm{M}}$	e_N	ε_N	R^2	*mare* (%)
• *Puromycin-treated data set,* $J=5$							
0	0	134.413223	0.015372	—	1.53e+5	0.5542	18.
1	5	152.072336	0.029454	0.114	0.343	0.7771	14.94
2	5	184.624253	0.044340	8.13e−6	6.97e−9	0.9330	9.198
3	5	**212.683743**	**0.064121**	3.09e−8	5.09e−11	0.9613	7.
• *EV data set,* $J=8$							
0	0	0.001957	−0.169663	—	0.0176	−0.8679	24.79
1	8	0.028507	0.301639	0.144	3.01e−5	0.0422	29.16
2	8	0.085433	1.314163	0.00445	3.76e−9	0.4085	27.51
3	5	**0.10564**3	**1.70269**0	1.05e−6	6.08e−15	0.9379	21.71

If the 2-BS (or in general the k-BS) process fails to converge for data sets with replicated observations we then propose a practical way, which is demonstrated in the following, for the computational improvement of the 2-BS process when it is applied on such data sets.

Any data set showing replications can be re-arranged in order to help the k-BS method to converge. With this re-arrangement of the Puromycin data set (which contains two V_i readings for each $C_{\mathrm{S},i}$ value), the 2-BS (or the 1-BS) process can now be applied adopting as initial vector $\hat{\boldsymbol{\theta}}_0$ the MM solution vector (using the first two observations) as we did in Example 3.1. Recall that this $\hat{\boldsymbol{\theta}}_0$ cannot be calculated in the case of the original (non-modified) Puromycin data set. The suggestion for helping the performance of the calculations is that we first split the Puromycin data set (and any data set that contains two readings for the depending variable for each value of the independent variable), say $\mathfrak{P}$, into two subsets, say $\mathfrak{P}_1$ and $\mathfrak{P}_2$. Each set contains the two "non-replicated" parts of the original data set and sorted in an increasing order of the C_{S} values, i.e.

$$\mathfrak{P}_1 := \big\{(C_{\mathrm{S},i}, V_i)\big\}_{i=1,3,5,\ldots,11} \quad \text{and} \quad \mathfrak{P}_2 := \big\{(C_{\mathrm{S},i}, V_i)\big\}_{i=2,4,6,\ldots,12} \,. \tag{3.10}$$

Note that $C_{\mathrm{S},i+2} > C_{\mathrm{S},i}$ for $i = 1, 3, ..., 9$ and $i = 2, 4, ..., 10$. In order to re-join them back into a single data set (of 12 observations), with a "smooth" transition from the (increasing) $C_{\mathrm{S},i}$ values of $\mathfrak{P}_1$ data set to the (also increasing) $C_{\mathrm{S},i}$ values of $\mathfrak{P}_2$, we adopt the $\mathfrak{P}_1$ data set as is, and then the observations of $\mathfrak{P}_2$ are put in the reversed (decreasing) order, i.e.

$$\mathfrak{P} = \big\{(C_{\mathrm{S},i}, V_i)\big\}_{i=1,3,5,\ldots,11,12,10,8,\ldots,2} \,. \tag{3.11}$$

The 2-BS process can then be applied, with convergence error threshold $e = 10^{-4}$, while we adopt the MM solution $(V_{\max}, K_{\mathrm{M}})$, as in Example 3.1, as the initial estimates' vector $\hat{\boldsymbol{\theta}}_0$. The first sub-Table of Table 4 provides the GN convergent J_N-th estimates $\hat{\boldsymbol{\theta}}_{N,J_N}$, that calculated with maximum number of GN iterations being only $J = 3$ (at every step $N = 1, 2, ..., 6$ of the 2-BS process). The resulted accuracy of the obtained NLLS estimate $(\hat{V}_{\max}, \hat{K}_{\mathrm{M}})$ is less than $10^{-5} < e$. Similarly, for the "replicated" EV data set, the accuracy for the NLLS estimates is less than $10^{-6} < e$; see the corresponding calculation on the second sub-Table of Table 4. The digits in bold represents again the accurate digits of the NLLS estimates.

Table 4: Convergence of the 2-BS processes for the MM model estimation of the re-arranged P and EV data sets.

N	J_N	$\hat{V}_{\max}$	$\hat{K}_{\mathrm{M}}$	e_N	ε_N	R^2	*mare* (%)
• *Puromycin-treated data set,* $J = 3$							
0	0	112.549618	0.009618	—	0.	1.	0.
1	1	112.549618	0.009618	0.	0.	1.	0.
2	3	170.999898	0.033674	5.76	2.07e+3	0.8736	10.56
3	3	210.839180	0.062539	0.7	75.7	0.9360	9.345
4	3	214.144962	0.064599	1.46e−5	4.73e−8	0.9480	7.739
5	3	212.825859	0.064726	7.6e−6	1.64e−8	0.9389	7.22
6	3	**212.683743**	**0.064121**	1.28e−6	7.47e−10	0.9613	7.
• *EV data set,* $J = 3$							
0	0	0.005853	−0.065438	—	2.33e−17	1.	0.
1	1	0.005853	−0.065438	2.6e−18	1.3e−18	1.	4.e−14
2	3	0.017440	0.158092	0.102	0.000647	0.6203	14.26
3	3	0.078120	1.285978	0.458	0.00152	0.9553	19.51
4	3	0.126137	2.411911	0.325	9.89e−6	0.9655	14.05
5	3	0.104970	1.694986	0.272	0.000761	0.9419	21.76
6	3	**0.105643**	**1.702690**	2.93e−7	2.52e−16	0.9379	21.71

The re-arrangement, as in (3.11), of the data set (which affects the order in which the observations are sequentially inserted into the BS process) can also improve the performance of the BS process even for initial guesses $\hat{\boldsymbol{\theta}}_0$ for which the 2-BS, or even the 1-BS, process normally could not converge. The following Example demonstrates this improvement.

Example 3.3. Considering the initial estimates' guess $\hat{\boldsymbol{\theta}}_0 = (80, 0)$ and letting $e = 10^{-4}$ to be the convergence error threshold, the 2-BS (as well as the 1-BS process) fails to converge, when it is applied on the original Puromycin data set. However, the 2-BS process converges, to the requested NLLS estimate, when the data set is re-arranged, as described in (3.11), even with few $(J = 3)$ allowed GN iterations at every step of 2-BS, or 1-BS, process. Table 5 presents the performance of the 2-BS approach (first sub-Table) as well as of the 1-BS approach

(second sub-Table). The accuracy of the obtained NLLS estimates derived from the 2-BS process is less than $10^{-5} < e$, while the 1-BS process results an accuracy less than $10^{-4} < e$. The digits in bold represents also here the accurate digits of the NLLS estimates.

Table 5: Convergence of the 1-BS and the 2-BS processes, applied on the re-arranged Puromycin data set.

N	J_N	$\hat{V}_{\max}$	$\hat{K}_{\mathrm{M}}$	e_N	ε_N	R^2	*mare* (%)
• *2-BS process,* $J = 3$							
0	0	80.	0.	—	1.33e+4	−0.3832	11.39
1	3	112.410518	0.009554	2.04	275.	1.	0.062
2	3	170.980951	0.033657	5.8	2.1e+3	0.8736	10.56
3	3	210.839081	0.062538	0.702	76.2	0.9360	9.345
4	3	214.144962	0.064599	1.46e−5	4.74e−8	0.9480	7.739
5	3	212.825859	0.064726	7.6e−6	1.64e−8	0.9389	7.22
6	3	**212.683743**	**0.064121**	1.28e−6	7.47e−10	0.9613	7.
• *1-BS process,* $J = 3$							
0	0	80.	0.	—	1.33e+4	−0.3832	11.39
1	3	112.410518	0.009554	2.04	275.	1.	0.062
2	3	136.051800	0.017418	0.563	25.3	0.8983	5.843
3	3	172.114585	0.034640	1.93	221.	0.8739	10.62
4	3	199.818136	0.053452	0.398	13.1	0.9129	10.46
5	3	210.857217	0.062575	0.00718	0.0089	0.9360	9.341
6	3	211.078437	0.062765	9.91e−10	8.75e−12	0.9471	8.015
7	3	214.144962	0.064599	9.91e−06	2.13e−8	0.9480	7.739
8	3	213.560104	0.067269	0.00066	0.000118	0.9397	7.37
9	3	212.825859	0.064726	0.000225	1.56e−5	0.9389	7.22
10	3	212.086616	0.062938	5.28e−5	1.07e−6	0.9440	6.93
11	3	**212.683743**	**0.064121**	1.02e−5	4.46e−8	0.9613	7.

4. DISCUSSION

Certain aspects of the MM model, so essential in Risk Analysis as far as to form an enzyme-substrate complex especially to pharmacokinetics studies, were discussed in this paper, either theoretical (see Theorem 2.1) or computational (see the provided examples in Section 3). As far as the optimal design is concerned, recall Kitsos [14] and (1.8), the design depends on the nonlinear term K_{M}. When the D-optimal design problem was viewed from the MM model perspective it can be formed into the following compact form:

If $C_{\mathrm{S}} \in (0, C_{\mathrm{U}}]$ the locally D-optimal design at $K_{\mathrm{M}} = K_0$ which allocates the half of the observations V with optimum concentration

$$C_{\mathrm{S}}^{\mathrm{opt}} = \frac{K_0 C_{\mathrm{U}}}{2K_0 + U}\,, \tag{4.1}$$

with C_{U} being the maximum allowable substrate concentration. If $C_{\mathrm{U}} \gg K_0$ the locally D-optimal design ξ is

$$\xi^* = \begin{pmatrix} C_{\mathrm{U}} & K_0 \\ 0.5 & 0.5 \end{pmatrix}. \tag{4.2}$$

The corresponding value of the determinant of the D-optimal design is

$$d = \frac{V_{\max}^2 C_{\mathrm{U}}^6}{16 K_0^2 (K_0 + C_{\mathrm{U}})^6}\ . \tag{4.3}$$

See also Endrenyi and Chan [6]. If $C_{\mathrm{S}} \in [C_{\mathrm{L}}, C_{\mathrm{U}}]$ then the optimum C_{S}, through (4.1), is given by

$$C_{\mathrm{S}}^{\mathrm{opt}} = \frac{K_0 C_{\mathrm{U}}}{2K_0 + (C_{\mathrm{U}} - C_{\mathrm{L}})}\ , \qquad K_0 > 0, \quad 0 < C_{\mathrm{L}} < C_{\mathrm{U}}\ . \tag{4.4}$$

From the above relations and the average-per-observation information matrix as in (1.8) it is clear, due to Theorem 2.1, that there might be more than one NLLS estimates. This is a new point of view of the design and actually a crucial one. One could choose as "best" among the D-optimal designs the one which provides minimum value of the corresponding (4.3) evaluation, which is a common situation. Therefore, there might be (locally) D-optimal designs corresponding to the analytical real-valued NLLS estimates. The final adopted D-optimal design can be chosen in principle, as noted also above, to be the one that provides minimum $\det\big(\mathbf{M}(\hat{\boldsymbol{\theta}}, \xi)\big)$. It is clear from relations (4.3) and (4.4) that the right choice of the existent different values for the parameter $\boldsymbol{\theta}$ is essential for the Risk Analysis study under investigation. It is why we provide static or sequential design approach to reach the appropriate selected real value for $\boldsymbol{\theta}$. It is therefore crucial what we prove: there is always one real value, and thus the Risk Analysis can always proceeded. How we proceed on Risk Analysis were more than one real value for $\boldsymbol{\theta}$ exists, has been extensively discussed, see Theorem 2.1 and the Examples.

An analytic methodology for the nonlinear least squares estimation (NLLS) was also introduced and compared against the four known linearization technics. The analytic formulation of this method indicated that the NLLS estimation of the MM model was, in general, not unique. Moreover, an iterative scheme for the NLLS estimation was also introduced, called the Batch Sequential (BS) process, and tested in various cases of data sets which showing readings replication or not. Despite that the BS is an iterative process, meaning that an initial estimates' guess is needed, a different approach was discussed and tested which provides an initial estimate that increases the convergence performance of the BS algorithm. Finally, certain examples demonstrate all the proposed methods.

ACKNOWLEDGMENTS

We would like to thank the referees for their useful comments which improve this paper.

REFERENCES

[1] Bates, D.M. and Watts, D.G. (1998). *Nonlinear-Regression Analysis and its Applications*, Oliver & Boyd, Edinburgh.

[2] Biebler, K.E.; Schreiber, K. and and Bode, R. (2009). *Solutions to the Michaelis–Menten kinetics are not necessarily unique*. In "AIP Conference Proceedings", 1148, 37 (doi: 10.1063/1.3225321).

[3] Currie, D.J. (1982). Estimating Michaelis–Menten parameters: bias, variance and experimental design, *Biometrics*, **38**, 907–919.

[4] Dowd, J.E. and Riggs, D.S. (1965). A comparison of estimates of Michaelis–Menten kinetic constants from various linear transformations, *J. of Biological Chemistry*, **240**, 863–869.

[5] Draper, N.R. and Smith, H. (1998). *Applied Regression Analysis, 3rd ed.*, Wiley.

[6] Endrenyi, L. and Chan, F.Y. (1981). Optimal design of experiments for the estimation of precise hyperbolic kinetic and binding parameters, *J. Theor. Biol.*, **90**, 241–263.

[7] Ford, I.; Titterington, D.M. and Kitsos, C.P. (1989). Recent advances in nonlinear experimental design, *Technometrics*, **31**(1), 49–60.

[8] Hadeler, K.; Jukić, D. and Sabo, K. (2007). Least squares problems for Michaelis–Menten kinetics, *Mathematical Methods in the Applied Sciences*, **30**(11), 1231–1241.

[9] Hill, P.D.H. (1980). D-optimal designs for partially nonlinear regression models, *Technometrics*, **22**, 275–276.

[10] Jennrich, R.J. (1969). Asymptotic properties of nonlinear least squares estimators, *Ann. Math. Stat.*, **40**, 633–643.

[11] Jukić, D.; Sabo, K. and Scitovski, R. (2007). Total least squares fitting Michaelis–Menten enzyme kinetic model function, *Journal of Computational and Applied Mathematics*, **201**, 230–246.

[12] Kitsos, C.P. (2013). *Optimal Experimental Design for Non-Linear Models*, Springer (ISSN: 2191-544X).

[13] Kitsos, C.P. (2012). *Cancer Bioassays: A Statistical Approach*, Lampert Acad. Pub.

[14] Kitsos, C.P. (2001). Design aspects for the Michaelis–Menten model, *Biometrical Letters*, **38**, 53–66.

[15] Kitsos, C.P. (1989). Fully sequential procedures in nonlinear design problems, *Comput. Stat. and Data Analysis*, **8**, 13–19.

[16] Kitsos, C.P. and Toulias, T.L. (2015). *On the Michaelis–Menten model.* In "Proceedings of the ICRA6 International Conference" (M. Guillen, A. Juan, H. Ramanhinho, I. Serra and C. Serrat, Eds.), 449–456.

[17] Michaelis, L. and Menten, M.L. (1913). Kinetics for intertase action, *Biochemische Zeitung*, **49**, 333–369.

[18] Toulias, T.L. and Kitsos, C.P. (2016). Fitting the Michaelis–Menten model, *Journal Of Computational And Applied Mathematics*, **296**, 303–319.