

Transform-both-sides nonlinear models for *in vitro* pharmacokinetic experiments

A. H. M. Mahbub Latif and Steven G. Gilmour

Institute of Statistical Research and Training (ISRT), University of Dhaka, Dhaka-1000, Bangladesh
and

Southampton Statistical Sciences Research Institute & Institute for Life Sciences
University of Southampton, Highfield, Southampton, SO17 1BJ, UK

June 27, 2014

Abstract

Transform-both-sides (TBS) nonlinear models have proved useful in many experimental applications including those in pharmaceutical sciences and biochemistry. The maximum likelihood (ML) method is commonly used to fit TBS nonlinear models, where the regression and transformation parameters are estimated simultaneously. In this paper, an analysis of variance (anova) based method is described in detail for estimating TBS nonlinear models from randomized experiments. It estimates the transformation parameter from the full treatment model and then the regression parameters are estimated conditionally on this estimate of the transformation parameter. The anova method is computationally simpler compared with the ML method of estimation and allows a more natural separation of different sources of lack of fit. Simulation studies show that the anova method can provide unbiased estimators of complex TBS nonlinear models, such as TBS random coefficient nonlinear regression (RCNLR) models and TBS fixed coefficient nonlinear regression (FCNLR) models with random block effects.

Keywords: Nonlinear mixed effects model, pure error and lack of fit, random block effects

1 Introduction

In many pharmacokinetic experiments, the main goal is to identify enzymes that are related to the metabolic process of the substrate of interest.¹ Such experiments are conducted at the very early stages of the drug development process. Since most of the enzymes that are involved in drug metabolism are located in the liver, human liver microsomes (HLMs) are used in these *in vitro* experiments.² Experimental runs are conducted for each HLM at different levels of substrate concentration and the response, the initial rate of reaction, is measured from each experimental run. The relationship between such a response and the substrate concentration is usually nonlinear and so it is assessed from the size of the nonlinear regression parameters.

In such pharmacokinetic experiments, the Michaelis-Menten model, with an additive normal error term, is widely assumed to model the effect of substrate concentration on the rate of reaction for each HLM. To accommodate HLM to HLM variation in the analysis, HLMs could be assumed to be either a fixed or a random treatment factor depending on the objective of the experiment and also on the study design considered. However, considering the HLM as a fixed treatment factor restricts the interpretation of the model parameters only to the HLMs used in the experiment, which would have very little use in pharmacokinetic studies. On the other hand, if HLM is considered as a random treatment factor in the model, it is often possible to make interpretations of the model parameters over a population of HLMs. In particular, it is assumed that variation in the nonlinear relationship between HLMs is random.

To model the rate of reaction as a function of random HLM and fixed substrate concentration, nonlinear mixed effects models can be used. The maximum likelihood (ML) method of estimation for a nonlinear mixed effects model depends on normal distributional assumptions regarding the random

effects and error terms. The classical assumptions regarding errors might not be satisfied in practice and a transformation of the response can be used in order to adjust a skewed and/or heteroskedastic error distribution to improve the approximation. For linear fixed effects models, the technique of transforming only the response has been widely used for many years either to stabilize the error variance or to simplify the regression function.³ Since most nonlinear regression models are derived from theoretical arguments, the regression parameters have scientific interpretations, which might be altered if only the response is transformed. Transforming both the response and regression function with the same transformation keeps the interpretation of the regression parameters the same as on the original scale.⁴ However, transforming both sides of the regression model increases the complexity of model fitting compared with transforming only the response. The transformations that are commonly used in statistical analysis may contain parameter(s), which either could be assumed as known or need to be estimated from the data at hand. A detailed discussion of the ML approach for fitting the transform-both-sides (TBS) fixed coefficient nonlinear regression (FCNLR) model can be found in Carroll and Ruppert's book⁵. Oberg and Davidian⁶ extended Carroll and Ruppert's approach to the analysis of TBS random coefficient nonlinear regression (RCNLR) models, see also Gilberg et al.⁷ Note that we are transforming both the response and the entire nonlinear regression function, and not just the regressors as in a fractional polynomial model.

Blocking, or local control, is one of the important tools of design of experiments, which is used to reduce the impact of experimental error on estimating treatment effects. Blocking can be useful in pharmacokinetic studies because there are situations where the experimental equipment might not allow the intended number of runs to be made at the same time, so that runs are made in batches. Usually estimating the block effects is not the main goal of the study, it is only used to compare treatments between homogeneous units. Ogliari and Andrade⁸ analyzed longitudinal data using nonlinear models in randomized block designs.

Though in this paper we discuss the analysis of pharmacokinetic experiments, the proposed method can also be used for other randomized experiments. In all published applications of TBS models, the corresponding transformation parameter is estimated by maximizing the likelihood function corresponding to the TBS nonlinear model - similar to the other parameters of the model. This approach does not use the fact that the transformation parameter only deals with the distribution of the error terms and has no connection to the nonlinear model under investigation. In this paper, we will discuss a method for estimating the transformation parameter by considering an appropriate full treatment model, which corresponds to the randomization used in the experiment. The remaining parameters of the nonlinear model are estimated conditionally after transforming both sides of the nonlinear model with the estimate of the transformation parameter. We will also discuss the advantages of this approach over estimating all the parameters simultaneously, using some simulation studies.

In Section 2 the TBS-FCNLR model is described and the usual ML estimation procedure is briefly discussed and the new analysis of variance (anova) based method is described. The anova method for the TBS-RCNLR model and the TBS-FCNLR model with random block effects are described in Section 3 and Section 4, respectively. A number of simulation studies are reported in Section 5 for comparing the ML and anova methods, and also for assessing the performance of the anova method in estimating the parameters of complex TBS nonlinear models.

2 Transform-both-sides nonlinear fixed effects models

The main goal of this paper is to develop a methodology for analysing enzyme kinetic and other TBS-RCNLR models for *in-vitro* pharmacokinetic studies using HLMs. In this section, we consider experiments on a single HLM and discuss the method for the TBS-FCNLR model, which will be extended for the mixed effects model in Section 3. Let n be the number of available experimental units and assume that we will use R experimental conditions (*treatments*) x_1, \dots, x_R . We also assume that at least two treatments have more than one replicate and let n_r be the number of experimental units assigned to treatment x_r , where $\sum_r n_r = n$. Assume that the design is completely randomized. Let $y_{i(r)}$ denote the response measured from the i^{th} experimental unit ($i = 1, \dots, n$), where the subscript " $i(r)$ " indicates that the treatment x_r was assigned to the i^{th} experimental unit. We use

this notation, following Hinkelmann et al.¹⁰, to emphasis the link between the design and the model, and allow easier extension to more complex designs. For notational convenience, we drop “ r ” from the subscript in the following sections, i.e. $y_i = y_{i(r)}$.

We assume the following nonlinear regression model for describing the relationship between the response and the treatment randomly assigned to the i^{th} experimental unit

$$\eta(y_i) = f(x_r; \boldsymbol{\theta}), \quad (1)$$

where $\eta(\cdot)$ is some measure of location, $\boldsymbol{\theta}$ is the p -dimensional vector of regression parameters to be estimated and the regression function $f(\cdot)$ is assumed to be nonlinear in at least one of the parameters $\boldsymbol{\theta}$. The least squares (LS) method of estimation is often used to fit a model of the type (1) if the classical assumptions (i.e. independent and identical distributions, constant variance, etc.) regarding the additive error terms ϵ' are satisfied, so that

$$y_i = f(x_r; \boldsymbol{\theta}) + \epsilon'_i. \quad (2)$$

Moreover, if the errors are normally distributed, the ML method of estimation can be used to estimate the parameters and, for the model of type (2), the LS and ML methods are equivalent. The theory of estimation and inference for nonlinear regression models are well developed and for a detailed theoretical treatment see Seber and Wild¹¹, for example. Most of the standard statistical software packages (e.g. R, SAS, Stata, GenStat) have user-friendly routines for fitting nonlinear fixed effects models and a comparison of the performance of the existing software in fitting nonlinear mixed effects model can be found in Plan *et al.*¹².

If the classical assumptions regarding the errors cannot be assumed to be satisfied, a transformation technique can be used with the expectation that the transformed data will provide more reliable inferences. After transforming both sides of the nonlinear regression model (1) and assuming an additive error on the appropriate scale, the resulting model becomes

$$h(y_i, \lambda) = h(f(x_r; \boldsymbol{\theta}), \lambda) + \epsilon_i, \quad (3)$$

where $h(y, \lambda)$ is a family of transformations of y and λ is known as the transformation parameter, which may be a scalar or vector. The model (3) is known as the TBS-FCNLR model and has been in the statistics literature for many years.⁴ After the transformation, it is expected that the errors (ϵ_i) are independent and identically normally distributed with zero mean and a constant variance σ^2 , i.e. they satisfy the classical assumptions regarding the errors. In this paper, the Box-Cox family of transformations³,

$$h(y, \lambda) = \begin{cases} (y^\lambda - 1)/\lambda & \text{if } \lambda \neq 0; \\ \log y & \text{if } \lambda = 0, \end{cases} \quad (4)$$

is used for the TBS nonlinear regression modelling, where $y > 0$. The main objectives of considering Box-Cox transformation in the analysis are to remove the skewness in response and also to remove the dependence of the standard deviation of response on its mean. The size of the error variance could be important because large error variance is required to remove the skewness, whereas Box-Cox transformation can remove the dependency of standard deviation of the response on its mean irrespective of the size of the error variance.⁹ Note that the methods described in the following sections can also be defined for other families of transformations.

2.1 The ML method of estimation

In this section, the commonly used ML method is briefly reviewed in the context of the TBS-FCNLR model (3). Assume that the transformed response $z_i = h(y_i, \lambda)$ follows a normal distribution with mean $h(f(x_r; \boldsymbol{\theta}), \lambda)$ and constant variance σ^2 . The probability density function (pdf) of the response on the original scale $y_i = h^{-1}(z_i, \lambda)$ can be obtained by multiplying the Jacobian $y_i^{\lambda-1}$ of the transformation $Z_i \rightarrow Y_i$ with the pdf of the transformed response $h(y_i, \lambda)$. The log-likelihood function of y_1, \dots, y_n can be expressed as

$$l_m(\boldsymbol{\theta}, \lambda) = -n \log \left[\frac{\hat{\sigma}(\boldsymbol{\theta}, \lambda)}{\tilde{y}^{\lambda-1}} \right] - \frac{n}{2}, \quad (5)$$

where $\tilde{y} = \exp(\sum_i \log y_i/n)$ and the maximum likelihood estimator (MLE) of σ^2 is $\hat{\sigma}^2(\boldsymbol{\theta}, \lambda) = (1/n) \sum_i (h(y_i, \lambda) - h(f_r, \lambda))^2$, where $f_r = f(x_r; \boldsymbol{\theta})$ is defined for notational simplicity. The MLEs $\hat{\boldsymbol{\theta}}_m$ and $\hat{\lambda}_m$ can be obtained by maximizing the log-likelihood function $l_m(\boldsymbol{\theta}, \lambda)$, i.e. $(\hat{\boldsymbol{\theta}}_m, \hat{\lambda}_m)' = \arg \max_{\boldsymbol{\theta}, \lambda} l_m(\boldsymbol{\theta}, \lambda)$, which is equivalent to minimizing

$$\frac{n \hat{\sigma}^2(\boldsymbol{\theta}, \lambda)}{\tilde{y}^{2(\lambda-1)}} = \sum_i \left(\frac{h(y_i, \lambda) - h(f_r, \lambda)}{\tilde{y}^{\lambda-1}} \right)^2 = \sum_i e_i^2(x_r, \boldsymbol{\theta}, \lambda), \quad (6)$$

for $\boldsymbol{\theta}$ and λ simultaneously, where $e_i(x_r, \boldsymbol{\theta}, \lambda)$ is known as the *pseudo-regression* function.⁵

It is convenient to use a regression routine available in statistical software (e.g. the `nls` function in R) instead of a general optimization routine in fitting a regression model because, for the latter, the user needs to write the function to be optimized. Since nonlinear regression routines in standard statistical software do not allow the response to depend on a parameter, the TBS-FCNLR model (3) cannot be fitted directly. To overcome this, Carroll and Ruppert⁵ suggested using a dummy response D_i with all of its values being zero and regress it on the *pseudo-regression* function $e_i(x_r, \boldsymbol{\theta}, \lambda)$, which is defined in equation (6), to obtain the MLEs $\hat{\boldsymbol{\theta}}_m$ and $\hat{\lambda}_m$. This approach cannot find the MLE of the error variance σ^2 , however, but the MLE $\hat{\sigma}^2$ can be obtained from $\hat{\sigma}^2(\hat{\boldsymbol{\theta}}_m, \hat{\lambda}_m)$. The corresponding restricted maximum likelihood (REML) estimator of σ^2 is $\tilde{\sigma}^2 = n \hat{\sigma}^2(\hat{\boldsymbol{\theta}}_m, \hat{\lambda}_m)/(n - p - 1)$, where $(n - p - 1)$ is the number of degrees of freedom associated with the residual sum of squares. A detailed discussion of the ML method for estimating the TBS Michaelis-Menten fixed effects model can be found in Ruppert et al.¹⁹

2.1.1 Estimates of standard errors

The standard errors of the parameter estimates of the TBS-FCNLR model (3) can be obtained in several ways and a good discussion of different methods can be found in Carroll and Ruppert⁵. Two methods are compared in this paper using a simulation study: (i) standard errors obtained from the pseudo-regression analysis (“ML(a)”) and (ii) standard errors obtained by a parametric bootstrap approach (“ML(boot)”). For the “ML(boot)” method, a random sample of residuals $\hat{\epsilon}_1^*, \dots, \hat{\epsilon}_n^*$ is drawn from $\mathcal{N}(0, \hat{\sigma}^2)$ and the model $z_i^* = h(f(x_r; \hat{\boldsymbol{\theta}}_m), \hat{\lambda}_m) + \hat{\epsilon}_i^*$ is used to generate response $y_i^* = h^{-1}(z_i^*, \hat{\lambda}_m)$ for the bootstrap sample. The ML estimates $\hat{\boldsymbol{\theta}}_m^*$ and $\hat{\lambda}_m^*$ can be obtained using y_i^* as the response. This procedure is repeated a large number of times to obtain the sampling distributions of the estimates $\hat{\boldsymbol{\theta}}_m$ and $\hat{\lambda}_m$ and, hence, the corresponding standard errors.

2.2 The anova method of estimation

The ML method, described in the previous section, simultaneously estimates both the regression and transformation parameters of the TBS-FCNLR model, which leads to more difficult convergence issues compared with fitting untransformed nonlinear models. Some of these are due to there being a ridge of near-maximal solutions with wrong values of $\boldsymbol{\theta}$ being compensated for by wrong values of λ . In this section a new method, which we call the anova method, for estimating the parameters of the TBS-FCNLR model (3) is described.

2.2.1 Estimation of the transformation parameter

The anova method separates the estimation of the transformation parameter from the specific nonlinear model assumed. The variance structure follows that given by the analysis of variance determined by the randomization of the experiment and the transformation parameter is estimated to make the additivity assumption required for analysis of variance plausible. Then the regression parameters are estimated given the estimated transformation parameter and the variance structure. The role of the transformation parameter λ in a TBS-FCNLR model (3) is to select an appropriate scale for the response so that the distribution of the transformed response has constant variance and is at least approximately normal.

To describe the anova method of estimation, assume the following full treatment model for the transformed response corresponding to the i^{th} experimental unit

$$h(y_i, \lambda) = \mu_r(x) + \delta_i, \quad (7)$$

where the i^{th} experimental unit is assigned to the r^{th} treatment, $\mu_r(x)$ is the mean function of the transformed response corresponding to the r^{th} treatment and the error term corresponding to the i^{th} experimental unit, δ_i , is assumed to be independent and normally distributed with mean 0 and constant variance σ_e^2 . The complete randomization procedure involved in the experiment leads to an analysis based on such a linear model - see Hinkelmann et al.¹⁰ for details. The TBS-FCNLR model (3) is a submodel of the full treatment model (7) with $\mu_r(x) = h(f_r, \lambda)$.

The full treatment model (7) depends only on treatment-unit additivity and normality, and not on the nonlinear model assumed for describing the relationship between the response and the treatments. The full treatment model is commonly written as

$$h(y_i, \lambda) = \mu + \tau_r + \delta_i, \quad (8)$$

where μ is the overall mean and τ_r is the fixed effect of the r^{th} treatment. Model (8) is linear in the parameters, which in this case have closed form estimates, depending only on the treatments and the transformed response. The goal of considering the full treatment model for the transformed response is to estimate the transformation parameter and the error variance (pure error) with minimal assumptions, i.e. the treatment parameters of model (8) are nuisance parameters at this stage.

Since the regression parameters involved in model (8) can be expressed as functions of the treatments and the response, the corresponding log-likelihood function is $l_b(\lambda)$, which is similar to (5) with $\hat{\sigma}^2(\theta, \lambda)$ replaced by

$$\hat{\sigma}_e^2(\lambda) = \frac{1}{n} \sum_i (h(y_i, \lambda) - \hat{\mu}_r(x))^2, \quad (9)$$

which is the MLE of σ_e^2 for a given λ , where $\hat{\mu}_r(x)$ is obtained directly from the one-way anova linear model (8), given λ . The MLE $\hat{\lambda}_a = \arg \max_{\lambda} l_b(\lambda)$ can also be obtained as $\hat{\lambda}_a = \arg \min_{\lambda} \{\hat{\sigma}_e(\lambda)/\tilde{y}^{\lambda-1}\}$. Since the error mean square or the likelihood function can be expressed as a function of λ only, a one-dimensional grid search can be used to obtain $\hat{\lambda}_a$. The REML estimator of σ_e^2 is $\tilde{\sigma}_e^2 = n \hat{\sigma}_e^2(\hat{\lambda}_a)/(n-R)$, where $(n-R)$ is the error degrees of freedom in this case. This method of estimating the transformation parameter is identical to that of Box and Cox³ described for analyzing linear models when transforming only the response.

2.2.2 Estimation of the regression parameters

The transformation parameter deals with selecting the appropriate scale for the response that leads to the distribution of the response having approximately constant variance and normality. Once the appropriate scale for the response is selected, the TBS-FCNLR model for the response measured on the scale defined by $\hat{\lambda}_a$ can be expressed as

$$h(y_i, \hat{\lambda}_a) = h(f(x_r, \theta), \hat{\lambda}_a) + \epsilon_i, \quad (10)$$

where θ is the parameter vector to be estimated and the error terms are assumed to be normally distributed with zero mean and a constant variance σ^2 . As Box and Cox³ suggested in the context of linear models, the transformation parameter is being treated as known in the TBS nonlinear model (10).

The log-likelihood function and error mean square for the model (10) are $l_m(\theta, \hat{\lambda}_a)$ and $\hat{\sigma}^2(\theta, \hat{\lambda}_a)$, respectively. The MLE $\hat{\theta}_a$ can be obtained by maximizing the log-likelihood function $l_m(\theta, \hat{\lambda}_a)$ or minimizing the error mean square $\hat{\sigma}^2(\theta, \hat{\lambda}_a)/\tilde{y}^{\hat{\lambda}_a-1}$. Model (10) can be fitted using a nonlinear regression routine, treating the transformation parameter as known. The definition of a pseudo-regression function is not required. Fitting model (10) is computationally simpler than fitting model (3) because the inclusion of the transformation parameter not only adds a dimension to the optimisation required but also in many cases gives a likelihood surface with ridge-like features.

2.2.3 Estimation of standard errors

For the anova method of estimating parameters of TBS-FCNLR model (3), the method described in Section 2.2.1 is used to estimate the transformation parameter λ . A number of methods are available in the literature for estimating the standard error of the MLE $\hat{\lambda}_a$,¹³ the simplest way being from the information matrix corresponding to the log-likelihood function $l_b(\lambda)$ evaluated at the MLE $\hat{\lambda}_a$. Atkinson and Lawrance¹⁴ discussed a method of calculating the standard error of $\hat{\lambda}_a$ using the full information matrix of the linear full treatment model, similar to (7), evaluated at the mean transformed response. The nonlinear regression routine of any standard statistical software can be used to fit the TBS-FCNLR model with known transformation parameter (10) and can provide reliable estimates of error variance and standard errors of the estimates of the regression parameters $\hat{\theta}_a$. We label this method as “anova(a)” for obtaining standard errors of $\hat{\lambda}_a$ and $\hat{\theta}_a$. The standard errors of $\hat{\theta}_a$ and $\hat{\lambda}_a$ can also be obtained using a bootstrap method, which is similar to that we described for the ML approach in Section 2.1.1. In this case, the random samples of residuals are generated from $\mathcal{N}(0, \hat{\sigma}_e^2)$ and, hence, the responses are obtained.

The anova method estimates the regression parameters conditionally, for a given value of the transformation parameter. On the other hand, the ML method estimates the regression and transformation parameters simultaneously. However, use of the estimate $\hat{\lambda}$ in making inference about the regression parameters remains an issue for both the methods, so that the interpretation of the regression parameters is conditional on $\hat{\lambda}$ in either case. In the case of linear models with transformed responses, Bikel and Doksum¹⁵ showed that treating the estimates of the transformation parameter as known can inflate the asymptotic variance of the linear regression parameter estimates compared with what standard linear theory provides. In response to this criticism, Hinkley and Runger¹⁶ stated that the results of Bikel and Doksum¹⁵ have little significance in practical applications even though they are mathematically correct. They argued that the regression parameters have no meaning without a particular scale and parameter estimates at different scales are not comparable. For the anova method, we follow Hinkley and Runger¹⁶ and fix the transformation parameter λ at some value for estimating the standard errors of regression parameter estimates, i.e. we are interested in $SE(\hat{\theta}|\hat{\lambda})$.

2.2.4 Separation of lack of fit from pure error

Given our assumptions, the size of the lack of fit compared with pure error can be used to examine the validity of the assumed model. If the model fits the data well, then it is expected that the lack of fit sum of squares will not be much larger than the pure error sum of squares. The methods described in Draper and Smith¹⁷ for separating lack of fit from pure error can be used with both the ML and anova methods in the context of the TBS nonlinear model (3). The ratio of the lack of fit and pure error mean squares is commonly used as the test statistic for examining the significance of the lack of fit. The test statistic follows a central F-distribution if there is no lack of fit - see Seber and Wild¹¹ for details.

A second method can be defined only for the anova method. The residual sum of squares of the full treatment model (7) corresponds to the pure error sum of squares, whereas the residual sum of squares of the TBS nonlinear model (10) corresponds to the sum of pure error and lack of fit sums of squares. The lack of fit sum of squares can be obtained by subtracting the residual sum of squares of the full treatment model (7) from the residual sum of squares of the TBS nonlinear model (10). Since the TBS-FCNLR (10) is a special case of the full treatment model (7), the difference of the likelihood functions corresponding to these two models can be used to assess the size of the lack of fit. Since the full treatment model is used to estimate the pure error, the anova method provides more natural definitions of pure error and lack of fit of the assumed nonlinear model form than the ML method for which the nonlinear predictor can be computed using an incorrect λ , i.e. using ML a wrong nonlinear functional model form can be compensated for by a wrong scaling of the responses. Thus the anova method gives less ambiguous model selection than the usual ML method.

2.3 Example: TBS-FCNLR model

As an example, the *Puromycin* data set¹⁸, is analyzed to compare the ML and anova methods, described in Sections 2.1 and 2.2, in fitting a TBS-FCNLR model. The data were obtained from an experiment in which the velocity of an enzymatic reaction was measured in counts/min² at 6 different substrate concentration levels with and without treating the enzyme by Puromycin. Two experiments are conducted at each of the 6 substrate concentrations. The goal of the original experiment was to examine whether the introduction of Puromycin affects the velocity of the reaction, which requires the analysis of both the treated and untreated data simultaneously. To compare the ML and anova methods in fitting TBS-FCNLR model, only the data from the runs treated by Puromycin are analyzed here and we assume that each measurement is from a different experimental unit (though this is unclear from the description given).

As in the original analysis of the Puromycin data, the Michaelis-Menten equation

$$f(x_r; \boldsymbol{\theta}) = \frac{V x_r}{K + x_r} \quad (11)$$

is assumed to model the relationship between the velocity and the substrate concentration, where $\boldsymbol{\theta} = (V, K)'$ and the parameters V and K represent the maximal rate of velocity and the substrate concentration corresponding to the half-maximal velocity, respectively.

The analyses of the Puromycin data using the Michaelis-Menten (M-M) model, described in Bates and Watts¹⁸, and the TBS Michaelis-Menten model are shown in Table 1. The large value of the F-statistic for lack of fit indicates that the ordinary Michaelis-Menten model does not fit the data well (p-value < .0001), whereas the TBS Michaelis-Menten model fits the data reasonably well (p-value > .2). The estimates of the parameters and the corresponding standard errors are similar for both the ML and anova methods of estimation.

Table 1: Analysis of Puromycin Data¹⁸ using Michaelis-Menten fixed effects model

Methods	Estimates and standard errors (SEs)						Lack of fit
	\hat{V}	SE(\hat{V})	\hat{K}	SE(\hat{K})	$\hat{\lambda}$	SE($\hat{\lambda}$)	F value
M-M	212.68	6.95	.064	.008	–	–	147.75
ML	218.03	4.74	.075	.009	2.08	.79	1.52
anova	217.77	4.53	.074	.008	1.99	.47	1.15

3 TBS models with random nonlinear regression coefficients

Nonlinear mixed effects models are widely used for analyzing repeated measures data. Davidian and Giltinan²⁰ discussed their estimation and inference, and Pinheiro and Bates²¹ discussed the computational details in the context of using the R package `nlme`²². The theory of nonlinear mixed effects models is based on normality assumptions about the errors and random effects of the model. Similar to the fixed effects models, violations of the assumptions regarding random effects can be overcome by transforming the data.²³ For linear mixed effects models, Gurka et al.²⁴ discussed the methods for analyzing models for transformation of only the response. Little has been published on the transformation of nonlinear mixed effects models and in the following, a method for analyzing TBS-RCNLR models is described.

3.1 The model

In this section, the TBS-RCNLR model is introduced in the context of the pharmacokinetic study described in Section 1. Suppose we have S HLMs (randomly selected from a population of HLMs) and R levels of substrate concentration. A subset of these $R \times S$ combinations of the HLMs and substrate concentrations constitute the treatments. Suppose the intention is to conduct n experimental runs and treatments are assigned to runs at random, so that the experimental unit is a run.

Let $y_i = y_{i(rs)}$ be the response measured from the i^{th} experimental unit ($i = 1, \dots, n$) which was randomly assigned the treatment that is a combination of the r^{th} substrate concentration level x_r and s^{th} HLM ($r = 1, \dots, R; s = 1, \dots, S$). As before, assume that at least two treatments are replicated and let n_{rs} be the number of replications of the treatment corresponding to the concentration level x_r and the s^{th} HLM, where $\sum_{r,s} n_{rs} = n$. The usual form of error terms cannot be assumed if the response is modeled on its original scale. We assume that errors and random effects are normally distributed after transformation.

Consider the following TBS-RCNLR model for the response corresponding to the i^{th} experimental unit

$$h(y_{i(rs)}, \lambda) = h(f(x_r; \boldsymbol{\theta}_s), \lambda) + \epsilon_i, \quad \boldsymbol{\theta}_s = \mathbf{A}_s \boldsymbol{\beta} + \mathbf{b}_s, \quad (12)$$

where $\boldsymbol{\theta}_s$ denotes the p -dimensional parameter vector corresponding to the s^{th} HLM, ϵ_i denotes the random error term, \mathbf{b}_s denotes a p -dimensional vector of random effects, \mathbf{A}_s is the $p \times q$ dimensional indicator variable matrix corresponding to the s^{th} HLM, and $\boldsymbol{\beta}$ denotes the q -dimensional vector of population fixed effects. It is assumed that $\mathbf{b}_s \sim \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma}(\boldsymbol{\alpha}))$, $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$, \mathbf{b}_s and ϵ_i are independent, and $\boldsymbol{\alpha}$ is the vector of parameters which completely specifies the variance-covariance matrix $\boldsymbol{\Sigma}$.

3.2 The ML method of estimation : TBS-RCNLR

In this section, we briefly review the ML method for fitting the TBS-RCNLR model (12). To write the likelihood function, we first divide all the responses y_1, \dots, y_n into S independent vectors $\mathbf{y}_1, \dots, \mathbf{y}_S$, where \mathbf{y}_s is the response vector of order $n_s (= \sum_r n_{rs})$ corresponding to the s^{th} HLM and $\sum_s n_s = n$. Let \mathbf{x}_s be the vector of concentration levels corresponding to the experimental units of the response vector \mathbf{y}_s . According to model (12), given the random effects \mathbf{b}_s , the transformed response $\mathbf{z}_s = h(\mathbf{y}_s, \lambda)$ follows an n_s -dimensional normal distribution with mean vector $\boldsymbol{\mu}_s = (h(f(x; \boldsymbol{\theta}_s), \lambda))_{x \in \mathbf{x}_s}$ and covariance matrix $\sigma^2 \mathbf{V}$, where \mathbf{V} can be expressed as a function of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and σ^2 .

The marginal density function of the transformed response vector \mathbf{z}_s can be obtained by integrating out the random effects term as

$$g_z(\mathbf{z}_s) = \int \cdots \int g_{z|b}(\mathbf{z}_s | \mathbf{b}_s) g_b(\mathbf{b}_s) d\mathbf{b}_{s1} \cdots d\mathbf{b}_{sp}, \quad (13)$$

where $g_{z|b}(\cdot)$ and $g_b(\cdot)$ are probability density functions corresponding to $\mathcal{N}_{n_s}(\boldsymbol{\mu}_s, \sigma^2 \mathbf{V})$ and $\mathcal{N}_p(\mathbf{0}, \boldsymbol{\sigma})$, respectively. The multi-dimensional integration is often nonlinear in the random effects and is very complicated to solve analytically. A number of approaches are available in the literature for approximating multi-dimensional integration of this type, two of the most popular methods being Taylor series approximation²⁵ and Laplace's approximation^{26;27}.

The marginal distribution of the response vector \mathbf{y}_s can be obtained by multiplying the Jacobian of the transformation $\mathbf{Z}_s \rightarrow \mathbf{Y}_s$ with the approximate expression of (13). In this case the likelihood function is $l(\boldsymbol{\beta}, \boldsymbol{\alpha}, \lambda) = \prod_{s=1}^S g_y(\mathbf{y}_s)$, where $g_y(\cdot)$ is the pdf of \mathbf{y}_s . The MLE of the population regression parameter $\boldsymbol{\beta}$, parameters involved in the random effects $\boldsymbol{\alpha}$, and the transformation parameter λ can be obtained simultaneously by optimizing $l(\boldsymbol{\beta}, \boldsymbol{\alpha}, \lambda)$.

Gilberg et al.⁷ suggested a ML method for fitting TBS-RCNLR model, which has two steps. In the first step, initial estimates of the regression and transformation parameters are obtained using a pseudo-regression model. In the second step the nonlinear model is approximated to obtain the corresponding linear model using the approach suggested by Lindstrom and Bates²⁵ and this step provides the estimate of the random effects terms. These two steps are iterated until the procedure converges. Oberg and Davidian⁶ suggested a method that uses the Laplace transformation for approximating the corresponding log-likelihood function of the transformed response. Both of these approaches are computationally very intensive.

3.3 The anova method of estimation : TBS-RCNLR

The anova method, described in Section 2.2 for the TBS-FCNLR model, can be extended to the TBS-RCNLR model (12) by considering an appropriate full treatment model. In the following sec-

tions, methods for estimating the transformation parameter λ , population regression parameter β and variance components are described.

3.3.1 Estimation of the transformation parameter

To estimate the transformation parameter λ , consider the following full treatment model for the transformed response corresponding to the i^{th} experimental unit

$$h(y_{i(rs)}, \lambda) = \mu_{rs}(x) + \delta_i,$$

where μ_{rs} is the mean of the transformed response corresponding to the treatment that is obtained from the combination of the substrate concentration level x_r and HLM s , the errors are assumed to be independent and $\delta_i \sim \mathcal{N}(0, \sigma_e^2)$. In this case, the mean function μ_{rs} depends on the levels of two factors, substrate concentration level and HLM, so we can parameterize the mean function μ_{rs} as

$$\mu_{rs}(x) = \mu + \tau_{rs}, \quad (14)$$

where μ is the overall mean and τ_{rs} is the effect of the treatment associated with the s^{th} HLM and r^{th} concentration level. The regression parameters of the full treatment model do not have any significance other than contributing to the estimation of transformation parameter λ . We treat all the factors of the full treatment model as if they are fixed and we use the model form in (14). The ML estimator $\hat{\lambda}_a$ can be obtained by maximizing the corresponding log-likelihood function, and the details of obtaining the MLE $\hat{\lambda}_a$ and the corresponding standard errors are described in Section 2.2.

3.3.2 Estimation of the regression parameters and variance components

Using the MLE $\hat{\lambda}_a$ of the transformation parameter in the TBS-RCNLR model (12), the resulting model becomes

$$h(y_i, \hat{\lambda}_a) = h(f(x_r; \theta_s), \hat{\lambda}_a) + \epsilon_i, \quad \theta_s = \mathbf{A}_s \beta + \mathbf{b}_s, \quad (15)$$

where the parameters to be estimated are the population regression parameters and variance components. The likelihood function corresponding to model (15) can be obtained as described in Section 3.2. In this case, the response vector corresponding to the s^{th} HLM $\mathbf{z}_s = h(\mathbf{y}_s, \hat{\lambda}_a)$ follows a multivariate normal distribution with mean vector $\boldsymbol{\mu}_s = (h(f(x; \theta_s), \hat{\lambda}_a))_{x \in \mathbf{x}_s}$ and covariance matrix $\sigma^2 \mathbf{V}$. The likelihood function is simpler than that of the model (12) because the transformation parameter is considered as known. Model (15) is similar to the standard nonlinear mixed effects model which can be fit using the methods described in the books of Davidian and Giltinan²⁰, and Pinheiro and Bates²¹ using standard software, e.g. the `nlme` function of R, although this might not be the best.¹² The anova method requires more than one replication of at least one of the treatments under investigation for estimating the transformation parameter.

3.3.3 Example: TBS-RCNLR

As an example of TBS-RCNLR model, the Phenacetin O-Deethylation (POD) data set, which was first analysed by Belle *et al.*¹, is used in this section. One of the objectives of their study was to identify cytochrome P450 (CYP) enzymes that catalyze POD. The CYP enzymes are important because they play a vital role in the metabolism of drugs. In the study, 19 HLMs and four concentration levels of POD are selected, and each combination of HLM and concentration level are replicated three times. Belle *et al.*¹ used the HLM-specific Michaelis-Menten equation $f(x, \theta_s) = V_s x / (K_s + x)$ to analyse the data in the original paper. The transform-both-sides Michaelis-Menten mixed effects model

$$h(y_{i(rs)}, \lambda) = h(V_s x_r / (K_s + x_r), \lambda) + \epsilon_i,$$

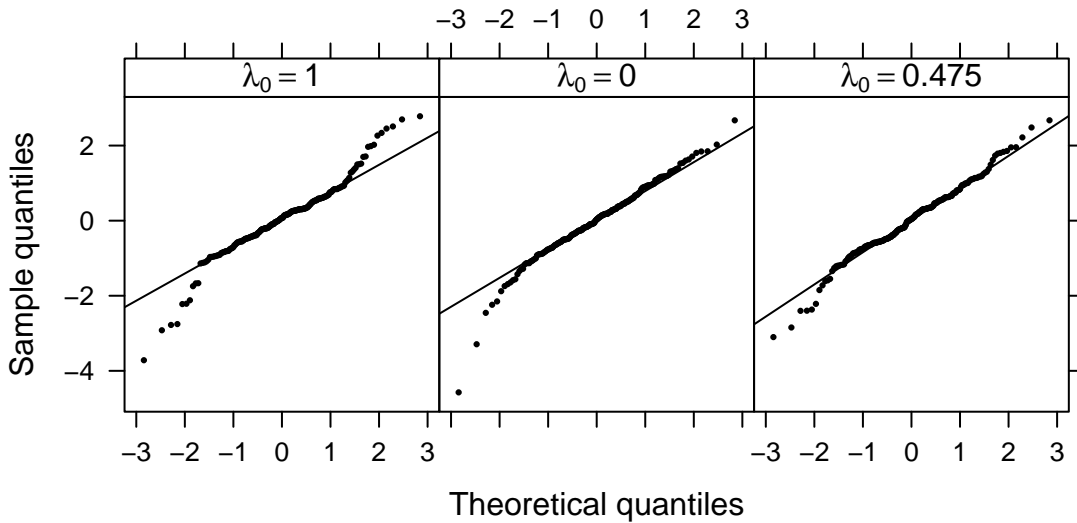
is considered for modelling the rate of reaction $y_{i(rs)}$ measured from the i^{th} experiment that corresponds to the s^{th} HLM and r^{th} level of concentration, where $\theta_s = (\log V_s, \log K_s)' = (V, K)' + \mathbf{b}_s$ and compared to the model defined in (7), $\mathbf{A}_s = \mathbf{I}_2$ and $\beta = (V, K)'$ are used. Assume that $\mathbf{b}_s \sim \mathcal{N}_2(\mathbf{0}, \Sigma)$, where $\Sigma = \text{diag}\{\sigma_V^2, \sigma_K^2\}$. Using the anova method of estimation, the estimate and the corresponding

standard error of the transformation parameter are $\hat{\lambda} = 0.475$ and $SE(\hat{\lambda}) = 0.054$, respectively. The estimates of the population regression parameters $(V, K)'$ and the random effects $(\sigma_V^2, \sigma_K^2)'$ given the transformation parameter $\lambda_0 \in \{1, 0, \hat{\lambda}\}$ are shown in Table 2. The estimates and corresponding standard errors are found to be similar for the three cases of Box-Cox transformation considered: $\lambda_0 = 1$ (no transformation), $\lambda_0 = 0$ (transform-both-sides by log transformation), and $\lambda_0 = \hat{\lambda} = 0.475$ (transform-both-sides by MLE of transformation parameter). Figure 1 shows a comparison between standardized residuals of the fitted TBS Michaelis-Menten mixed effects models with transformation parameters 1, 0, and 0.475, respectively. The Q-Q normal plot shows that the standardized residuals of the model corresponding to the MLE of the transformation parameter (i.e. $\hat{\lambda} = 0.475$) is closer to the standard normal distribution than that of obtained from the TBS models with no transformation (i.e. $\lambda = 1$) and log-transformation (i.e. $\lambda = 0$).

Table 2: Analysis of Phenacetin O-Deethylation data¹ using ordinary and transform-both-sides Michaelis-Menten mixed effects models.

Models	\hat{V}	$SE(\hat{V})$	\hat{K}	$SE(\hat{K})$	$\hat{\sigma}_V$	$\hat{\sigma}_K$	λ_0
Model-1	6.991	0.116	3.192	0.189	0.498	0.778	1.000
Model-2	6.915	0.128	2.899	0.157	0.556	0.661	0.000
Model-3	6.946	0.123	3.006	0.168	0.532	0.709	0.475

Figure 1: Q-Q normal plot of standardized residuals obtained from transform-both-sides Michaelis-Menten mixed effects models with transformation parameter $\lambda_0 \in \{1, 0, 0.475\}$.



4 TBS nonlinear models with random block effects

In this section the methods of analyzing TBS nonlinear models are discussed when the experiments are conducted in blocks.

4.1 TBS-FCNLR models

The method of analyzing TBS-FCNLR model, discussed in Sections 2.1 and 2.2, is now extended by incorporating blocking in the experimental setup. Consider an experimental scenario, in which pre-specified groups of treatments x_1, \dots, x_R are randomly allocated to blocks and, within each block, treatments are randomly allocated to runs. Assume that there are J blocks and each block has b runs.

If $b < R$ then the design is called an incomplete block design, which is most often used in applications. In this case, the total number of experimental units is $n = bJ$.

Let $y_{ji(r)}$ be the response corresponding to the i^{th} experimental unit in the j^{th} block, which was randomly allocated to the r^{th} treatment. Consider the model for the transformed response

$$h(y_{ji(r)}, \lambda) = h(f(x_r; \boldsymbol{\theta}), \lambda) + B_j + \epsilon_{ji}, \quad (16)$$

where ϵ_{ji} is the random error term corresponding to the i^{th} experimental unit in the j^{th} block and B_j is the random block effect corresponding to the j^{th} block. It is assumed that $\epsilon_{ji} \sim \mathcal{N}(0, \sigma^2)$, $B_j \sim \mathcal{N}(0, \sigma_b^2)$ and ϵ_{ij} and B_j are independent. The goal is to estimate the regression parameter $\boldsymbol{\theta}$ and the variance components σ^2 and σ_b^2 . To our knowledge, analysis of models of the type (16) has not been discussed in the literature in the context of nonlinear models.

4.1.1 The anova method of estimation

As in previous sections, a full treatment model is assumed in the anova method for estimating the transformation parameter, i.e.

$$h(y_{ji(r)}, \lambda) = \mu_r(x) + \beta_j + \delta_{ji}, \quad (17)$$

where $\mu_r(x)$ denotes the mean function corresponding to the r^{th} concentration level, β_j denotes the fixed effect of the j^{th} block, and random error terms δ_{ji} are assumed to be normally distributed with zero mean and constant variance σ_e^2 . For a given value of λ , the analysis of model (17) follows the usual analysis for an incomplete block design.^{10;28;29} The MLE $\hat{\lambda}_a$ is the estimator that maximizes the likelihood function corresponding to model (17) and hence the estimate of the pure error mean square $\hat{\sigma}_e^2$ can also be obtained.

Using the MLE $\hat{\lambda}_a$, the TBS nonlinear regression model with random block effects (16) becomes

$$h(y_{ji(r)}, \hat{\lambda}_a) = h(f(x_r; \boldsymbol{\theta}), \hat{\lambda}_a) + B_j + \epsilon_{ji}, \quad (18)$$

where the regression parameters $\boldsymbol{\theta}$, and variance components σ_b^2 and σ^2 are the parameters of interest. To write down the likelihood function, we use the fact that responses within each block are correlated and responses between blocks are independent. Let $\mathbf{y}_j = (y_{j1}, \dots, y_{jb})'$ be the vector of responses corresponding to the j^{th} block and assume that the \mathbf{y}_j 's are independent, $j = 1, \dots, J$. Following the assumptions of the model (16), $h(\mathbf{y}_j, \lambda) \sim \mathcal{N}_b(h(\boldsymbol{\mu}_j, \hat{\lambda}_a), \sigma^2 \Sigma_j)$, where $\Sigma_j = \mathbf{I}_b(\sigma^2 + \sigma_b^2)/\sigma^2 + \mathbf{J}_b(\sigma_b^2/\sigma^2)$, \mathbf{I}_b is an identity matrix of order b , \mathbf{J}_b is a b -dimensional square matrix of 1's, $\boldsymbol{\mu}_j = (f(x_{1j}; \boldsymbol{\theta}), \dots, f(x_{bj}; \boldsymbol{\theta}))'$ and $\{x_{1j}, \dots, x_{bj}\}$ is the set of concentration levels assigned to b plots of the j^{th} block.

The MLEs of the regression parameters and variance components can be obtained by maximizing the likelihood function corresponding to model (18). The generalized nonlinear least squares method of estimation can be used. Models of type (18) can be fitted using statistical software, e.g. the `gnls` function in R.

5 Simulation studies

In this section, we discuss the results of some simulation studies for examining the performance of the methods described in Sections 2–4. For the TBS-FCNLR model, described in Sections 2.1–2.2, we will show the comparison between the ML method and the anova method. The performance of the anova method is reported for fitting the TBS-RCNLR model and TBS-FCNLR model with random block effects.

5.1 TBS-FCNLR model

The Michaelis-Menten equation (11) is used in the simulation study. The response corresponding to the i^{th} experimental unit, which is assigned to the concentration level x_r , is modeled as

$$y_{i(r)} = h^{-1}(h(f(x_r; \boldsymbol{\theta}), \lambda) + \epsilon_i), \quad (19)$$

where $f(x_r, \theta) = Vx_r/(K + x_r)$, errors are assumed to be independent and $\epsilon \sim \mathcal{N}(0, \sigma^2)$. We used $V = 5$, $K = 3$, and $\sigma = .1$ for simulating responses from model (19). Simulation results are examined for different values of the true transformation parameter $\lambda \in [-1, 2]$ and also for different sets of concentration levels, i.e. different designs.

Table 3 reports the average bias and standard deviation (SD) of the estimates \hat{V} , \hat{K} , $\hat{\lambda}$, and $\hat{\sigma}$, computed from 2000 simulated samples. In this case, 10 replications of the design $X_1 = \{1, 3, 5, 7, 9\}$ are considered in the simulation, i.e. the total number of experimental units used is $n = 50$. The results show that all the parameter estimates are almost unbiased in the sense that the bias is found to be no more than 1/10 of the corresponding standard deviation and we will use this definition of unbiasedness in interpreting the remainder of the simulation study. We note that the bias is higher for $\lambda = -1$ as was also found by Ruppert *et al.*¹⁹ The variability of the estimates \hat{V} and \hat{K} tend to decrease as the true value of the transformation parameter increases. On the other hand, the variability of the estimates $\hat{\lambda}$ and $\hat{\sigma}$ do not depend on the true value of λ . On comparing the anova and ML methods, the standard deviations are found to be equal up to two decimal places for each λ value. The percentage of simulations which converged is also given and the anova method is found to be more stable in this regard compared with the ML method, especially when the true transformation parameter is around zero. Similar results are obtained from the simulations with 5 replications of the design X_1 , but these are not shown. This shows that the separation of the estimation of the transformation parameter from the estimation of the nonlinear location parameters does indeed have the expected benefit of improving the stability of estimation.

Table 3: Bias and standard deviation (SD) of the Michaelis-Menten fixed effects model parameters for the ML and anova methods of estimation. 2000 simulations are used with true parameter values $V = 5$, $K = 3$, $\sigma = 0.1$, and $\lambda \in [-1, 2]$. The design $X_1 = (1, 3, 5, 7, 9)$ and $n = 50$ are used.

λ	Method	Bias (SD)				%Converged
		\hat{V}	\hat{K}	$\hat{\lambda}$	$\hat{\sigma}$	
2.0	ML	.0004 (.0301)	.0008 (.0561)	-.0163 (.2685)	.0014 (.0277)	100
	anova	.0005 (.0302)	.0009 (.0563)	-.0326 (.2743)	-.0011 (.0275)	100
1.5	ML	.0009 (.0496)	.0018 (.0882)	-.0256 (.2726)	.0006 (.0278)	100
	anova	.0010 (.0497)	.0019 (.0884)	-.0325 (.2743)	-.0011 (.0275)	100
1.0	ML	.0017 (.0783)	.0033 (.1302)	-.0362 (.2746)	-.0004 (.0279)	100
	anova	.0017 (.0783)	.0033 (.1301)	-.0322 (.2742)	-.0010 (.0276)	100
0.5	ML	.0028 (.1201)	.0051 (.1797)	-.0343 (.2694)	-.0003 (.0274)	97
	anova	.0026 (.1203)	.0047 (.1799)	-.0314 (.2734)	-.0010 (.0277)	100
0.0	ML	.0034 (.1883)	.0064 (.2424)	-.0587 (.2838)	-.0023 (.0285)	91
	anova	.0041 (.1876)	.0063 (.2414)	-.0293 (.2711)	-.0008 (.0277)	99
-0.5	ML	.0062 (.3079)	.0086 (.3341)	-.0509 (.2555)	-.0021 (.0263)	99
	anova	.0095 (.3084)	.0105 (.3348)	-.0193 (.2585)	-.0001 (.0270)	100
-1.0	ML	.0288 (.5395)	.0243 (.5073)	.0059 (.2074)	.0020 (.0244)	93
	anova	.0398 (.5420)	.0324 (.5094)	.0292 (.2095)	.0035 (.0248)	93

The pseudo-regression based and bootstrap methods of obtaining standard errors are described in Section 2.1.1 and Section 2.2.3 for the ML and the anova method, respectively. The comparisons between these are reported in Table 4, where the estimates are computed from 2000 simulations using model (19). For the bootstrap method, 500 bootstrap samples are generated from each simulated sample to obtain the standard errors of the estimates. The same values of the regression parameters, error variance, and the design are considered in this case but results are reported only for $\lambda \in \{0.5, 1, 1.5\}$ because the bootstrap method requires more computational time. The methods of estimating standard errors are compared using the ratio of the average standard error (AvgSE) to the standard deviation (SD) of the estimates, $RE = \text{AvgSE}/\text{SD}$. For a method that correctly estimates the standard error, the RE value will be close to unity. The results show that the bootstrap approach outperforms the pseudo-regression based method in most cases for both the ML and anova methods. On average, a large number of experimental units and a small value of the transformation parameter lead to more

reliable estimates of the standard errors. If the pseudo-regression based method is considered, the ML method outperforms the anova method for the estimates \hat{V} and \hat{K} , but the anova method performs better than the ML method for the estimates $\hat{\lambda}$. On average, the ML and the anova methods perform equally well if the bootstrap approach is used for estimating the standard errors of the estimates.

Table 4: Ratio of the average estimated SE to SD (RE=AvgSE/SD) of the Michaelis-Menten fixed effects model parameters for the ML and anova methods of estimation. 2,000 simulations are used with true parameter values $V = 5$, $K = 3$, $\sigma = 0.1$, and $\lambda = \{1.5, 1.0, 0.5\}$. The design $X_1 = \{1, 3, 5, 7, 9\}$ and $n = \{25, 50\}$ are used.

n	λ	Method	RE(\hat{V})		RE(\hat{K})		RE($\hat{\lambda}$)	
			(a)	(boot)	(a)	(boot)	(a)	(boot)
25	1.5	ML	0.962	0.984	0.942	0.965	1.317	1.014
		anova	0.910	1.006	0.886	0.987	0.847	1.021
	1.0	ML	0.982	1.011	0.964	0.996	1.291	1.013
		anova	0.937	1.020	0.916	1.009	0.837	1.004
	0.5	ML	0.965	0.996	0.914	0.950	1.232	0.981
		anova	0.939	1.009	0.895	0.976	0.826	0.991
50	1.5	ML	1.004	1.014	0.997	1.008	1.375	1.009
		anova	0.985	1.027	0.976	1.021	0.933	1.014
	1.0	ML	0.995	1.006	0.976	0.989	1.361	1.001
		anova	0.980	1.019	0.959	1.002	0.920	1.003
	0.5	ML	0.981	0.992	0.970	0.983	1.388	1.026
		anova	0.964	0.996	0.954	0.994	0.923	1.006

5.2 TBS-RCNLR model

As an example of the TBS-RCNLR model, we simulate the pharmacokinetic study described at the beginning of Section 3.1. The relationship between the treatment corresponding to the s^{th} HLM and substrate concentration x_r , is modelled as

$$f(x_r; \boldsymbol{\theta}_s) = \frac{V_s x_r}{K_s + x_r}, \quad \boldsymbol{\theta}_s = (\log V_s, \log K_s)' = (V, K)' + \mathbf{b}_s,$$

where $\boldsymbol{\theta}_s$ and \mathbf{b}_s are the parameter vector and random effects, respectively, corresponding to the s^{th} HLM. It is assumed that $\mathbf{b}_s \sim \mathcal{N}_2(\mathbf{0}, \Sigma)$, where $\Sigma = \text{diag}\{\sigma_V^2, \sigma_K^2\}$. This model is similar to the nonlinear model considered in Section 3.3.3.

The response corresponding to the i^{th} experimental unit, which is associated with the s^{th} HLM and concentration level x_r , is generated from

$$y_{i(rs)} = h^{-1}(h(f(x_r; \boldsymbol{\theta}_s), \lambda) + \epsilon_i),$$

where error terms $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$. In this case $\lambda = -1$ does not give higher biases than other values, although we do not know why. The parameter values used in the simulations are $V = K = 1$, $\sigma_V = \sigma_K = 0.1$, $\sigma = 0.02$, and $\lambda \in [-1, 2]$.

Table 5 shows the average bias and standard deviation of the estimates of the regression parameters, variance components and transformation parameter from 2000 simulations. In this case, 10 HLMs and substrate concentration levels $X_1 = \{1, 3, 5, 7, 9\}$ are used to simulate the data and each treatment is replicated 10 times, i.e. 500 experimental units are used in the simulation. The bias and standard deviation corresponding to $\hat{\lambda}$ are similar to the anova method for the TBS-FCNLR model (see Table 3), because in both cases the treatment model is considered for estimating the parameter. All the parameter estimates are almost unbiased. The pattern of the estimates of regression parameters and variance components remains the same for different values of the transformation parameter.

The effects of using different numbers of HLMs, replications, and concentration levels for the estimate of bias and SD are summarized in Table 6. In this case, all the combinations of the designs

Table 5: Bias and SD of the parameter estimates of TBS Michaelis-Menten mixed effects model. True values of the parameters used for the simulation are $V = 1$, $K = 1$, $\sigma_V = 0.1$, $\sigma_K = 0.1$, and $\sigma = 0.02$. The design $X_1 = \{1, 3, 5, 7, 9\}$ and 10 HLMs are used with 10 replications.

λ	Bias (SD)				
	\hat{V}	\hat{K}	$\hat{\sigma}_V$	$\hat{\sigma}_K$	$\hat{\lambda}$
2.0	.0067 (.0326)	.0028 (.0368)	-.0073 (.0234)	-.0076 (.0292)	-.0120 (.1399)
1.5	.0066 (.0325)	.0037 (.0348)	-.0072 (.0233)	-.0079 (.0275)	-.0113 (.1418)
1.0	.0066 (.0325)	.0044 (.0336)	-.0072 (.0232)	-.0075 (.0261)	-.0111 (.1425)
0.5	.0066 (.0325)	.0049 (.0330)	-.0071 (.0233)	-.0069 (.0249)	-.0110 (.1428)
0.0	.0067 (.0325)	.0053 (.0326)	-.0071 (.0233)	-.0067 (.0244)	-.0110 (.1431)
-0.5	.0067 (.0324)	.0053 (.0324)	-.0071 (.0233)	-.0065 (.0241)	-.0110 (.1430)
-1.0	.0067 (.0324)	.0054 (.0323)	-.0071 (.0232)	-.0064 (.0239)	-.0110 (.1430)

$\{X_1, X_2, X_3\}$, the number of HLMs $S \in \{5, 10\}$ and the number of replications $\{5, 10\}$ are used, where $X_2 = \{1, 3, 6, 9\}$ and $X_3 = \{1, 5, 9\}$, and the results based on 2000 simulations are shown only for $\lambda = 0.5$. All the estimates are found to be almost unbiased with respect to the size of the corresponding standard deviations. Increasing the number of HLMs reduces the variability of the estimates of the regression parameters and random effects more than if the number of replications is increased. The accuracy of the estimate $\hat{\lambda}$ depends on the number of experimental units per treatment used in the experiment, large numbers of experimental units provide more accurate estimates. In simulation, we found that increasing the number of HLMs improves the accuracy of the $SD(\hat{V})$ and $SD(\hat{K})$, but increasing the number of replications only improves the accuracy of $SD(\hat{\lambda})$.

Table 6: Bias and SD of the parameter estimates of TBS Michaelis-Menten mixed effects model. True values of the parameters used for the simulation are $V = 1$, $K = 1$, $\sigma_V = 0.1$, $\sigma_K = 0.1$, $\sigma = 0.02$, and $\lambda = 0.5$. Three designs $X_1 = \{1, 3, 5, 7, 9\}$, $X_2 = \{1, 3, 6, 9\}$ and $X_3 = \{1, 5, 9\}$ are used for the simulations.

Design	S/n_s	Bias (SD)				
		\hat{V}	\hat{K}	$\hat{\sigma}_V$	$\hat{\sigma}_K$	$\hat{\lambda}$
X_1	5/5	.0071 (.0452)	.0019 (.0488)	-.0138 (.0320)	-.0129 (.0350)	-.0251 (.3106)
	5/10	.0044 (.0457)	.0076 (.0466)	-.0162 (.0314)	-.0153 (.0331)	-.0141 (.2055)
	10/5	.0049 (.0331)	.0033 (.0343)	-.0065 (.0235)	-.0082 (.0268)	-.0158 (.2160)
	10/10	.0066 (.0325)	.0049 (.0330)	-.0071 (.0233)	-.0069 (.0249)	-.0110 (.1428)
X_2	5/5	.0043 (.0447)	.0032 (.0494)	-.0145 (.0311)	-.0123 (.0350)	-.0280 (.3171)
	5/10	.0063 (.0464)	.0041 (.0454)	-.0139 (.0330)	-.0146 (.0332)	-.0140 (.2115)
	10/5	.0035 (.0324)	.0026 (.0341)	-.0067 (.0229)	-.0081 (.0258)	-.0074 (.2234)
	10/10	.0045 (.0317)	.0046 (.0331)	-.0067 (.0227)	-.0085 (.0245)	-.0066 (.1461)
X_3	5/5	.0052 (.0457)	.0029 (.0492)	-.0152 (.0317)	-.0121 (.0359)	-.0245 (.3398)
	5/10	.0055 (.0455)	.0024 (.0457)	-.0142 (.0318)	-.0153 (.0339)	-.0149 (.2169)
	10/5	.0044 (.0321)	.0031 (.0348)	-.0063 (.0230)	-.0078 (.0266)	-.0148 (.2421)
	10/10	.0045 (.0315)	.0042 (.0329)	-.0068 (.0237)	-.0085 (.0246)	-.0055 (.1556)

5.3 TBS-FCNLR model with random block effects

The response corresponding to the i^{th} experimental unit, in the j^{th} block, which is assigned the r^{th} concentration level, can be generated from the model

$$y_{ji(r)} = h^{-1}(h(f(x_r; \boldsymbol{\theta}), \lambda) + B_j + \epsilon_{ji}), \quad (20)$$

where, similar to (11), the Michaelis-Menten equation $f(x; \boldsymbol{\theta}) = Vx/(K + x)$ is considered as the nonlinear function, the random block effect is $B_j \sim \mathcal{N}(0, \sigma_b^2)$ and error terms are $\epsilon_{ji} \sim \mathcal{N}(0, \sigma^2)$. We assume that B_j and ϵ_{ji} are independent. For the simulation, we have considered a balanced

incomplete block design in which 30 experimental units are arranged in 10 blocks and each block has 3 runs. True values of the parameters used in the simulation are $V = 5$, $K = 3$, $\sigma_b = 0.8$, $\sigma = 0.02$, and $X_1 = \{1, 3, 5, 7, 9\}$.

Table 7 shows the average bias, SD and ratio of the average SE to SD of the parameter estimates, computed from 2000 simulations. All the estimates are found to be almost unbiased with respect to the size of the corresponding standard deviation. The size of the variability does not depend on the true value of λ for the estimates $\hat{\sigma}_b$ and $\hat{\lambda}$.

Table 7: Bias, SD and ratio of average SE to SD of the parameter estimates of the TBS Michaelis-Menten fixed effects model with random blocks. True values of the parameters used in the simulation are $V = 5$, $K = 3$, $\sigma = .02$, $\sigma_b = .08$. The design $X_1 = \{1, 3, 5, 7, 9\}$ and 10 blocks each has 3 plots are used.

λ	Bias (SD) [RE]			
	\hat{V}	\hat{K}	$\hat{\sigma}_b$	$\hat{\lambda}$
2.0	-.0001 (.0106) [.898]	-.0005 (.0276) [.900]	.0043 (.0259)	-.0011 (.1832) [.674]
1.5	.0007 (.0172) [.871]	.0016 (.0473) [.874]	.0026 (.0271)	-.0047 (.1910) [.644]
1.0	.0006 (.0259) [.925]	.0019 (.0720) [.863]	.0001 (.0253)	-.0113 (.1885) [.650]
0.5	.0012 (.0560) [.960]	.0004 (.0829) [.874]	-.0014 (.0254)	-.0135 (.1961) [.622]
0.0	.0051 (.1420) [.954]	.0008 (.0792) [.969]	-.0020 (.0243)	-.0133 (.1858) [.659]
-0.5	.0121 (.3235) [.939]	.0048 (.1709) [.948]	-.0020 (.0241)	-.0080 (.1828) [.665]
-1.0	.0746 (.7164) [.913]	.0459 (.4522) [.910]	-.0021 (.0232)	-.0068 (.1710) [.679]

6 Conclusion

The results show that both the ML and anova methods estimate the regression and transformation parameters with very small bias compared with the corresponding standard deviations. For the TBS nonlinear model, accurate estimation of the standard errors of regression parameter estimates is an issue, but the bootstrap method can estimate the standard error more precisely compared with the pseudo-regression based approach, for both the ML and anova methods. The anova method is found to be more stable compared to the ML method in some situations.

The anova method is not only computationally simpler than the ML method, but also can provide unbiased estimators of the parameters. Estimates of the regression parameters depend on the number of factor levels used, whereas the estimation of the transformation parameter depends on the number of replications used in the experiment.

Estimation of pharmacokinetic models is often problematic and one of the reasons is that we often do not make most use of the benefits implied by the structure of a well-designed experiment. Indeed, the standard analysis of experimental data is identical to that which would be performed were the data obtained from an observational study. The methods described here are very easy to use and have the potential to give experimenters more reliable conclusions from their data analysis. Of course, the methods described here can only be applied to replicated experiments. However, in such cases, they perform very well and we recommend them for practical use.

Acknowledgement

This work was supported by the Engineering and Physical Sciences Research Council (UK) grant EP/C541715/1: Unifying Approaches to Design of Experiments. Part of this work was carried out at the Isaac Newton Institute in 2011, as part of the Design and Analysis of Experiments programme. We thank two anonymous referees for many helpful comments.

References

- [1] Belle DJ, Ring BJ, Allerheiligen SRB, Heathman MA, O'Brien LM, Sinha V, et al. A population approach to enzyme characterization and identification: application to phenacetin O-deethylation. *Pharmaceutical Research*. 2000;17(12):1531–1536.
- [2] Gunaratna C. Drug metabolism & pharmacokinetics in drug discovery: a primer for bioanalytical chemists, part I. *Current Separations*. 2000;19(1):17–24.
- [3] Box G, Cox D. An analysis of transformations. *Journal of the Royal Statistical Society, Series B*. 1964;26:211–252.
- [4] Snee R. An alternative approach to fitting models when re-expression of the response is useful. *Journal of Quality Technology*. 1986;18(4):211–225.
- [5] Carroll RJ, Ruppert D. Transformation and Weighting in Regression. Chapman and Hall/CRC; 1988.
- [6] Oberg A, Davidian M. Estimating data transformations in nonlinear mixed effects models. *Biometrics*. 2000;56(1):65–72.
- [7] Gilberg F, Urfer W, Edler L. Heteroscedastic nonlinear regression models with random effects and their application to enzyme kinetic data. *Biometrical Journal*. 1999;41(5):543–557.
- [8] Ogliari PJ, Andrade DF. Analysing longitudinal data via nonlinear models in randomized block designs. *Computational Statistics and Data Analysis*. 2001;36(3):319–332.
- [9] Carroll RJ, Ruppert D. Power transformations when fitting theoretical models to data. *Journal of the American Statistical Association*. 1984;79(386):321–328.
- [10] Hinkelmann K, Kempthorne O. Design and Analysis of Experiments: Introduction to Experimental Designs. 2nd ed. Wiley; 2008.
- [11] Seber GAF, Wild CJ. Nonlinear Regression. Wiley-IEEE; 1989.
- [12] Plan EL, Maloney A, Mentré F, Karlsson MO, Bertrand J. Performance comparison of various maximum likelihood nonlinear mixed-effects estimation methods for dose-response models. *The AAPS Journal*. 2013;14(3):320–432.
- [13] Sakia RM. The Box-Cox transformation technique: a review. *The Statistician*. 1992;41:169–178.
- [14] Atkinson AC, Lawrance AJ. A comparison of asymptotically equivalent test statistics for regression transformation. *Biometrika*. 1989;76(2):223–229.
- [15] Bickel PJ, Doksum KA. An analysis of transformations revisited. *Journal of the American Statistical Association*. 1981;76(374):296–311.
- [16] Hinkley DV, Runger G. The analysis of transformed data. *Journal of the American Statistical Association*. 1984;79(386):302–309.
- [17] Draper NR, Smith H. Applied Regression Analysis. John Wiley & sons; 1998.
- [18] Bates D, Watts D. Nonlinear Regression Analysis and its Applications. John Wiley & sons; 1988.
- [19] Ruppert D, Cressie N, Carroll RJ. A transformation/weighting model for estimating Michaelis-Menten parameters. *Biometrics*. 1989;45:637–656.
- [20] Davidian M, Giltinan D. Nonlinear Models for Repeated Measurement Data. Chapman and Hall, New York; 1995.
- [21] Pinheiro JC, Bates DM. Mixed-Effects Models in S and S-PLUS. Springer Verlag; 2000.

- [22] Pinheiro J, Bates D, DebRoy S, Sarkar D, et al. nlme: Linear and Nonlinear Mixed Effects Models; 2008. R package version 3.1-90.
- [23] Solomon P. Transformations for components of variance and covariance. *Biometrika*. 1985; 72(2): 233–239.
- [24] Gurka MJ, Edwards LJ, Muller KE, Kupper LL. Extending the Box-Cox transformation to the linear mixed model. *Journal of the Royal Statistical Society, Series A*. 2006;169(2):273–288.
- [25] Lindstrom MJ, Bates DM. Nonlinear mixed effects models for repeated measures data. *Biometrics*. 1990;46(3):673–687.
- [26] Wolfinger R. Laplace’s approximation for nonlinear mixed models. *Biometrika*. 1993;80(4):791–795.
- [27] Vonesh EF. A note on the use of Laplace’s approximation for nonlinear mixed-effects models. *Biometrika*. 1996;83(2):447–452.
- [28] Bailey RA. Design of Comparative Experiments. Cambridge University Press; 2008.
- [29] Mead R, Gilmour SG, Mead A. Statistical Principles for the Design of Experiments. Cambridge University Press; 2012.