An update on modeling dose–response relationships: Accounting for correlated data structure and heterogeneous error variance in linear and nonlinear mixed models[1,2]

M. A. D. Gonçalves,* N. M. Bello,†[3] S. S. Dritz,*
M. D. Tokach,‡ J. M. DeRouchey,‡ J. C. Woodworth,‡ and R. D. Goodband‡

*Department of Diagnostic Medicine/Pathobiology, College of Veterinary Medicine;
†Department of Statistics, College of Arts and Sciences; and ‡Department of Animal Sciences and Industry,
College of Agriculture, Kansas State University, Manhattan 66506-0201

ABSTRACT: Advanced methods for dose–response assessments are used to estimate the minimum concentrations of a nutrient that maximizes a given outcome of interest, thereby determining nutritional requirements for optimal performance. Contrary to standard modeling assumptions, experimental data often present a design structure that includes correlations between observations (i.e., blocking, nesting, etc.) as well as heterogeneity of error variances; either can mislead inference if disregarded. Our objective is to demonstrate practical implementation of linear and nonlinear mixed models for dose–response relationships accounting for correlated data structure and heterogeneous error variances. To illustrate, we modeled data from a randomized complete block design study to evaluate the standardized ileal digestible (SID) Trp:Lys ratio dose–response on G:F of nursery pigs. A base linear mixed model was fitted to explore the functional form of G:F relative to Trp:Lys ratios and assess model assumptions. Next, we fitted 3 competing dose–response mixed models to G:F, namely a quadratic polynomial (QP) model, a broken-line linear (BLL) ascending model, and a broken-line quadratic (BLQ) ascending model, all of which included heteroskedastic specifications, as dictated by the base model. The GLIMMIX procedure of SAS (version 9.4) was used to fit the base and QP models and the NLMIXED procedure was used to fit the BLL and BLQ models. We further illustrated the use of a grid search of initial parameter values to facilitate convergence and parameter estimation in nonlinear mixed models. Fit between competing dose–response models was compared using a maximum likelihood–based Bayesian information criterion (BIC). The QP, BLL, and BLQ models fitted on G:F of nursery pigs yielded BIC values of 353.7, 343.4, and 345.2, respectively, thus indicating a better fit of the BLL model. The BLL breakpoint estimate of the SID Trp:Lys ratio was 16.5% (95% confidence interval [16.1, 17.0]). Problems with the estimation process rendered results from the BLQ model questionable. Importantly, accounting for heterogeneous variance enhanced inferential precision as the breadth of the confidence interval for the mean breakpoint decreased by approximately 44%. In summary, the article illustrates the use of linear and nonlinear mixed models for dose–response relationships accounting for heterogeneous residual variances, discusses important diagnostics and their implications for inference, and provides practical recommendations for computational troubleshooting.

Key words: computational troubleshooting, dose–response, heterogeneous variances, linear and nonlinear mixed models

INTRODUCTION

Dose–response models are used to estimate the minimal concentrations of a nutrient that maximize a given outcome, thereby determining nutritional requirements for optimal performance. Polynomials and broken lines are functional forms commonly used in regression

models to estimate nutrient dose–response relationships (Robbins et al., 1979; Vedenov and Pesti, 2007; Pesti et al., 2009). These models are often used in a fixed-effects modeling framework that assumes mutually independent observations with homogeneously dispersed errors. Yet experimental data often present a design structure that includes correlations between observations (i.e., blocking, nesting, etc.) and heteroskedastic errors (Wiggans and VanRaden, 1991; Wolfinger, 1996). In fact, heterogeneity of residual variances, also known as heteroskedasticity, seems to be a relatively common phenomenon in animal production systems (Cernicchiaro et al., 2013; Gonçalves et al., 2015) and it can mislead inference if disregarded (Wiggans and VanRaden, 1991; Wolfinger, 1996).

Mixed models are particularly well suited to handling correlated data (Littell et al., 2006). Implementation of mixed models, however, is not without challenges, including convergence of the iterative estimation process, particularly when fitting nonlinear mixed models. A common problem is that models either fail to converge or converge to suboptimal solutions (i.e., local vs. global maxima). A grid search approach can assist the estimation process by providing initial parameter values over the likelihood surface and guiding the iterative process away from suboptimal solutions and facilitating a more efficient search for optimal solutions (Kiernan et al., 2012).

The main objective of this paper is to demonstrate practical implementation of linear and nonlinear mixed models methodology for dose–response relationships, accounting for correlated data structures and heterogeneous variances. Second, we illustrate techniques to facilitate computational implementation of these models.

MATERIALS AND METHODS

Data

We used the data set previously presented by Gonçalves et al. (2015) on G:F of nursery pigs fed experimental diets consisting of increasing levels of standardized ileal digestible (**SID**) Trp:Lys ratio. Briefly, data were collected under a randomized complete block design whereby 1,088 pigs arranged in pens of 24 to 27 pigs were blocked by average initial BW and randomly assigned to experimental diets (6 pens/diet) consisting of SID Trp:Lys ratios of 14.5, 16.5, 18.0, 19.5, 21.0, 22.5, and 24.5%. The response variable G:F in its observed scale ranged from 520 to 610 g BW gain/kg feed intake. Our choice of units for G:F as grams per kilogram feed intake rather than the more commonly used kilograms per kilogram was intended to ensure numerical stability in the estimation process, particularly for variance components. It is well described that when

```
data GF;
     input Block Trt PenID$ y  varGrp$;

     datalines;
1    0.145    184    562    Var2
1    0.165    187    593    Var1
1    0.180    182    579    Var1
1    0.195    189    610    Var2
1    0.210    186    583    Var2
1    0.225    188    604    Var2
1    0.245    176    572    Var2
2    0.145    145    549    Var2
2    0.165    185    574    Var1
2    0.180    162    592    Var1
2    0.195    164    586    Var2
2    0.210    183    602    Var2
2    0.225    181    593    Var2
2    0.245    180    594    Var2
3    0.145    168    520    Var2
3    0.165    165    587    Var1
3    0.180    166    590    Var1
3    0.195    147    582    Var2
3    0.210    167    606    Var2
3    0.225    174    588    Var2
3    0.245    173    578    Var2
4    0.145    154    563    Var2
4    0.165    170    588    Var1
4    0.180    169    577    Var1
4    0.195    175    555    Var2
4    0.210    150    566    Var2
4    0.225    146    563    Var2
4    0.245    148    551    Var2
5    0.145    172    529    Var2
5    0.165    149    571    Var1
5    0.180    152    575    Var1
5    0.195    161    568    Var2
5    0.210    159    593    Var2
5    0.225    171    573    Var2
5    0.245    153    604    Var2
6    0.145    156    535    Var2
6    0.165    151    579    Var1
6    0.180    163    581    Var1
6    0.195    155    568    Var2
6    0.210    157    591    Var2
6    0.225    160    585    Var2
```

Figure 1. Example data set from Gonçalves et al. (2015), which evaluated the effects of standardized ileal digestible (SID) Trp:Lys ratio on nursery pig performance with 2 residual variance groups (Var1 for [16.5 and 18.0% SID Trp:Lys ratios] and [14.5, 19.5, 21.0, 22.5, and 24.5% SID Trp:Lys ratios] for Var2).

the scale of the data is very small and close to internal tolerances of computational algorithms, convergence can be impaired and estimation suffers (Kiernan et al., 2012). Proper choice of unit scale is important to ensure computational stability in the estimation process but does not affect the nature of the conclusions. The raw data are reported in Fig. 1 as a SAS software (version 9.4; SAS Inst. Inc., Cary NC) data step and includes BW blocks (i.e., "Block"), SID Trp:Lys treatment (i.e.,

```
proc glimmix data=GF plots=studentpanel;
     class block trt;
     model y = trt / ddfm=kr;
     random intercept / subject = block;
     output out=igausout pred=p student=std;
     nloptions tech=nrridg;
     lsmeans trt / cl plot=meanplot(cl join);
run;
```

Figure 2A. Base model assuming homogeneous residual variance.

```
proc glimmix data=GF plots=studentpanel;
     class block trt varGrp;
     model y = trt / ddfm=kr;
     random intercept / subject = block;
     random _residual_ / group = varGrp;
     output out=igausout pred=p student=std;
     nloptions tech=nrridg;
     lsmeans trt / cl plot=meanplot(cl join);
run;
```

Figure 2B. Base model allowing for heterogeneous residual variances.

Figure 2. A) Base model assuming homogeneous residual variance. B) Base model allowing for heterogeneous residual variances.

"Trt"), pen identification (i.e., "PenID"), and the response G:F expressed in grams per kilogram.

Base Mixed Model: Specification and Implementation

We started by fitting a "base" linear mixed model to 1) explore possible functional forms of the relationship between G:F and SID Trp:Lys ratios; 2) evaluate model assumptions, in particular homogeneity of residual variances; and 3) obtain preliminary estimates of variance components (i.e., block and residual variances) that could later be used as starting values in dose–response linear and nonlinear mixed models. The base mixed model was specified as follows:

$$y_{ij} = \eta + \alpha_i + b_j + e_{ij}, \qquad [1]$$

in which y_{ij} is the G:F expressed in grams per kilogram associated with the experimental unit in block j assigned to SID Trp:Lys ratio i and η corresponds to an intercept, whereas α_i represents the differential effect of Trp:Lys ratio i (treated as a categorical variable); in turn, b_j is the random effect of the jth block with $b_j \sim N(0, \sigma_b^2)$ and e_{ij} is a random residual associated with the experimental unit in the jth block that received the ith SID Trp:Lys ratio whereby $e_{ij} \sim N(0, \sigma_e^2)$, and b_j and e_{ij} are assumed to be independent of each other.

The base mixed model was fitted with the GLIMMIX procedure of SAS software (version 9.4;

Fig. 2A) using its default estimation method REML. The Kenward–Roger's procedure was used to estimate degrees of freedom and adjust estimated SE for bias correction (Littell et al., 2006). To assess model assumptions, we plotted studentized residuals as a function of levels of the treatment factor (Fig. 3A). All observations had values of studentized residuals between −3 and 3, thereby indicating no evidence for any extreme observations beyond probabilistic expectation. However, it was apparent from Fig. 3A that the amount of dispersion of studentized residuals around their expected value 0 (i.e., horizontal like in Fig. 3A) was noticeably uneven across treatments. Note, for instance, that at an 18% SID Trp:Lys ratio, residuals were tightly clustered around 0, whereas for diets consisting of a 14.5 or 24.5% SID Trp:Lys ratio, residuals seemed to have the greatest dispersion around 0, thereby questioning the assumption of a homogeneous residual variance in the base mixed model.

Specification of Heterogeneous Residual Variances. To further evaluate potential heteroskedasticity, we expanded our base mixed model in Eq. [1] to accommodate heterogeneous residual variances such that $e_{ij} \sim N\left(0, \sigma_{e_{ij(k)}}^2\right)$ with subscripts indicating the kth level of a variance group to which the ijth observation corresponded. Figure 2B shows SAS software code to implement heteroskedastic specifications of the base mixed model. We defined alternative variance groups

consisting of 2, 3, 4, or 7 levels empirically defined from Fig. 3A, as treatment combinations having seemingly comparable residual dispersion. These variance groups are listed in Table 1 and are presented for illustration, realizing that this is not an exhaustive list. A commented SAS software code is available as a supplementary file (see the online version of the article at http://journalofanimalscience.org) to illustrate how the variance groups were defined. The choice of model with the best fitting heterogeneous variance specification was based on the Bayesian information criterion (**BIC;** Schwarz, 1978). Please refer to the model selection section below for more details.

Dose–Response Estimation Models: Specification and Implementation

Next, we considered 3 competing linear and nonlinear dose–response mixed models, namely a quadratic polynomial (**QP**) model, a broken-line linear (**BLL**) ascending model, and a broken-line quadratic (**BLQ**) ascending model. These competing models represent 3 commonly used functional forms of the relationship between nutrient requirement and the response G:F, based on the AA nutrition literature (Robbins et al., 2006; Pesti et al., 2009). The competing models were specified as follows:

Quadratic polynomial model:

$$y_{ij} = \beta_0 + \beta_{1,\text{QP}} X_i + \beta_{2,\text{QP}} X_i^2 + b_j + e_{ij}. \qquad [2]$$

Broken-line linear ascending model:

$$y_{ij} = \varphi_{\text{BLL}} + \beta_{\text{BLL}} \times (\omega_{\text{BLL}} - X_i) + b_j + e_{ij}$$
$$for\ X_i < \omega_{\text{BLL}}, \qquad [3]$$
and
$$y_{ij} = \varphi_{\text{BLL}} + b_j + e_{ij} \qquad for\ X_i \geq \omega_{\text{BLL}}.$$

Broken-line quadratic ascending model:

$$y_{ij} = \varphi_{\text{BLQ}} + \beta_{1,\,\text{BLQ}} \times (\omega_{\text{BLQ}} - X_i) + \beta_{2,\,\text{BLQ}} \times$$
$$(\omega_{\text{BLQ}} - X_i)^2 + b_j + e_{ij} \qquad for\ X_i < \omega_{\text{BLQ}}, \qquad [4]$$
and
$$y_{ij} = \varphi_{\text{BLQ}} + b_j + e_{ij} \qquad for\ X_i \geq \omega_{\text{BLQ}}.$$

In these models, y_{ij} is the observed G:F associated with the pen randomly assigned to the SID Trp:Lys ratio i within block j and X_i indicates the ith known SID Trp:Lys ratio. For all models, b_j is the random effect of the jth block with $b_j \sim N(0, \sigma_b^2)$ and e_{ij} is a random error associated with the experimental unit in the jth block that received the ith SID Trp:Lys ratio whereby $e_{ij} \sim N\left(0, \sigma_{e_{ij(k)}}^2\right)$, whereby the group composition of each the kth levels was defined by the best-fitting heteroskedastic base model (as described in previous section). Also, b_j and e_{ij} are as-
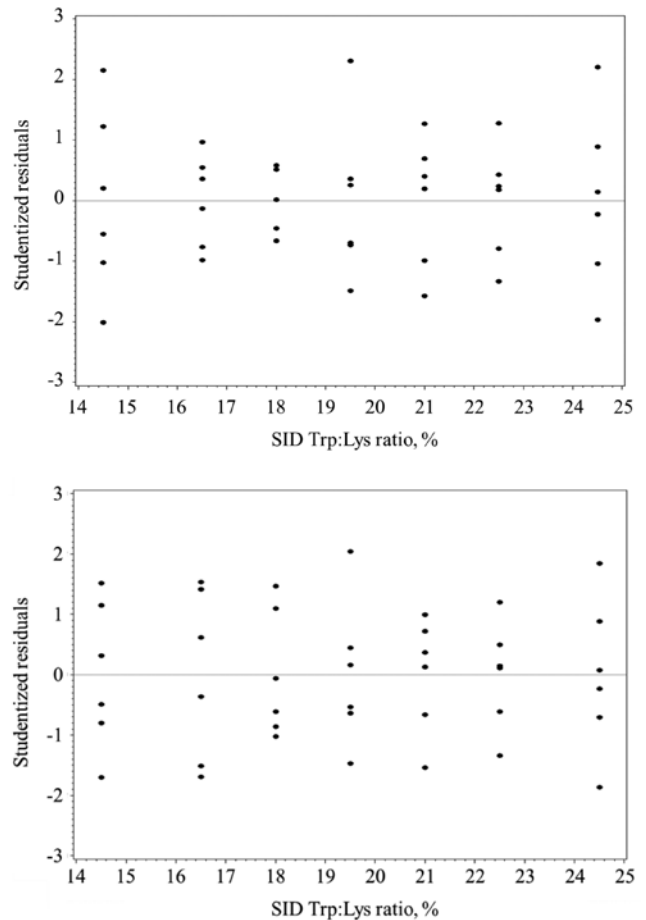

Figure 3. A) Studentized residuals of G:F, expressed in grams per kilogram, by treatment levels obtained from a base model fitted with a common residual variance. B) Studentized residuals of G:F, expressed in grams per kilogram, by treatment levels obtained from a base model allowing for heterogeneous of residual variances with 2 group variances ([16.5 and 18.0% standardized ileal digestible {SID} Trp:Lys ratios] vs. [14.5, 19.5, 21.0, 22.5, and 24.5% SID Trp:Lys ratios]).

sumed to be independent of each other. For the QP model, β_0 is the intercept and $\beta_{1,\text{QP}}$ and $\beta_{2,\text{QP}}$ are the corresponding unknown linear and quadratic regression coefficients relating X_i to the observed response y_{ij}. For the nonlinear models, φ_{BLL} and φ_{BLQ} indicate the unknown maximum response (i.e., plateau) under the BLL and BLQ models, respectively; in turn, β_{BLL}, $\beta_{1,\text{BLQ}}$, and $\beta_{2,\text{BLQ}}$ are the corresponding unknown regression coefficients describing the relationship between X_i and y_{ij} for values of X_i smaller than the plateau. Finally, ω_{BLL} and ω_{BLQ} are the unknown minimum levels of the SID Trp:Lys ratio to reach the plateau under the BLL and BLQ models, respectively. We note that our implementation of nonlinear mixed models takes into consideration the standard recommendation of the hierarchy principle for model building, whereby if a higher-order polynomial term is retained in the model, then the related lower-order terms are also kept in the model regardless of whether or not the coefficients for the lower-order terms are significant (Kutner et al., 2005).

Table 1. Bayesian information criterion (BIC) fit statistics for base models fitted assuming a homogeneous (i.e., common) variance or heterogeneous variances for 2, 3, 4, or 7 groups

Variance component	BIC	Group constituency
Common variance	303.6	No groups
2 variance groups – combination I	299.6	(16.5 and 18.0) vs. (14.5, 19.5, 21.0, 22.5, and 24.5)
2 variance groups – combination II	300.6	(14.5, 19.5, and 24.5) vs. (16.5, 18.0, 21.0, and 22.5)
3 variance groups	300.3	(14.5, 19.5, and 24.5) vs. (21.0 and 22.5) vs. (16.5 and 18.0)
4 variance groups	301.1	(14.5, 19.5, and 24.5) vs. (21.0 and 22.5) vs. (16.5) vs. (18.0)
7 variance groups	306.4	One group per treatment

The SAS software (version 9.4) was used to implement all dose–response models. The GLIMMIX procedure was used to fit the QP model whereas NLMIXED procedure was used for the nonlinear mixed models (i.e., BLL and BLQ). Under both procedures, the method of estimation was specified to be maximum likelihood (**ML**) to enable comparison of competing models. Figure 4 shows code used to implement the dose–response models using SAS software, in particular the QP model (Fig. 4A), the BLL model (Fig. 4B), and the BLQ model (Fig. 4C).

We note that, for the data set used in this article, the final dose–response models used for inference needed to account for heterogeneous residual variances, as dictated by preliminary data exploration using the base mixed model (see the previous section). We further illustrate the impact that disregarding heterogeneity of variances when modeling can have on the point estimates of the dose–response breakpoint as well as on its inference. Specifically, we compare estimates of nutrient requirements and respective inference based on the best fitting dose–response model with and without specification of heterogeneous residual variances.

Improving Computational Performance of Nonlinear Mixed Models. To facilitate the iterative estimation process of fitting nonlinear mixed models, we provided initial values for key parameters using the PARMS statement (Fig. 4B and 4C). Initial values for parameters φ_{BLL}, φ_{BLQ}, β_{BLL}, $\beta_{1,BLQ}$, $\beta_{2,BLQ}$, ω_{BLL}, and ω_{BLQ} were approximated from empirical scatter plots of the data and also from fitted values from the base mixed model. In turn, initial values for σ_b^2 and $\sigma_{e_k}^2$ were elicited using estimates from the base mixed model. For each parameter, at least 3 initial values were used to conduct a grid search so that for the BLL model, a total of 35 sets of starting values were evaluated, whereas at least 36 sets of starting values were considered for the BLQ model.

Estimating Confidence Intervals for Optimal Nutrient Requirements. For the nonlinear mixed models, the estimated mean breakpoint parameters (i.e., ω_{BLL} for the BLL model and ω_{BLQ} for the BLQ model) and their corresponding asymptotic confidence interval (**CI**; i.e., 95%) follow from ML estimation of model parameters; these values can be obtained directly from the SAS NLMIXED output. For the QP model, the estimated mean dose level at which the maximum response occurred is not automatically output and needs to be computed, as follows. We first obtained the first derivative of the fitted regression equation with respect to the predictor variable X (Pesti et al., 2009) and then equated the derivative to 0 and solved for X, thus obtaining the value of X that maximized the average response; this value was derived to be $-(\hat{\beta}_1)/(2\hat{\beta}_2)$. A $(1 − \alpha)$% CI for the estimated mean dose level using the QP model can be approximated using a graphical approach (Lavagnini and Magno, 2007). Briefly, the fitted QP equation is plotted over the dose levels with the desired estimated CI (i.e., 95% CI). Then, the maximum estimated response is projected on the y-axis using a horizontal line. The points of intersection of this horizontal line with the CI boundaries on the predicted line are then projected onto the x-axis as CI estimators of the optimum dose level.

An alternative approach to computing CI on the mean dose level at which the maximum response occurred involves numerical estimation through Monte Carlo simulation. Specifically, one may draw a large number of paired samples of parameters β_1 and β_2 from a multivariate normal distribution with mean parameter given by the point parameter estimates and variance–covariance values specified as the corresponding estimates of the sampling distributions of β_1 and β_2. Then, for each pair of β_1 and β_2 draws, one may compute the value $-(\beta_1)/(2\beta_2)$, thus obtaining an empirical distribution of the value of X that maximizes the average response under the QP model. The $(\alpha/2)$th and $(1 − \alpha/2)$th percentiles of this empirical distribution can then be used to report the $(1 − \alpha)$% CI for the estimated mean dose level that maximizes G:F based on the QP model.

Model Comparison. Competing mixed models used to evaluate 1) heterogeneity of error variances and 2) the functional form of the dose–response relationship (i.e., QP, BLL, and BLQ) were compared based on model fit using the BIC (Schwarz, 1978). It is noted that the competing models evaluated here either for heterogeneity of error variances or for functional form

```
proc glimmix data=GF method=MSPL ;
    class  block varGrp;
    model y = trt trt*trt / solution covb ;
    random intercept / subject=block;
    random _residual_ / group=varGrp;
    output out=igausout pred=p resid=r;
    nloptions tech=nrridg;
run;
```

Figure 4A. Quadratic polynomial mixed model with heterogeneous variance.

```
proc nlmixed data=GF maxiter=1000 gconv=0  start;
    bounds .145<R_BLL<.245;
    parms L_BLL= 578 582 586 U_L= -975 -1950 -3900 R_BLL= 0.15 0.16 0.17
    Block_Var= 6 11 22 vareVar1= 28 56 112 vareVar2= 134 268 536;
    z=(trt<R_BLL)*(R_BLL-trt);   * Characterize the model as non-linear;
    s2e = (varGrp="Var1") * vareVar1 + (varGrp="Var2") * vareVar2;
    model y ~ normal(L_BLL + U_L *(z) + beff, s2e);
    random beff ~ normal(0,Block_Var) subject=block out=blups;
    predict L_BLL + U_L*(z) out=ppp;
run;
```

Figure 4B. Broken-line linear mixed model with heterogeneous variance.

```
proc nlmixed data=GF maxiter=1000 gconv=0  start;
    bounds .145<R_BLQ<.245;
    parms L_BLQ= 578 582 586  U_Q1= -975 -1950 -3900
    U_Q2= -4685 -9369 -18738 -40000 -80000 R_BLQ=0.15 0.16 0.17
    Block_Var= 6 11 22 vareVar1= 28 56 112 vareVar2= 134 268 536;
    z=(trt<R_BLQ)*(R_BLQ-trt);
    s2e = (varGrp="Var1") * vareVar1 + (varGrp="Var2") * vareVar2;
    model y ~ normal(L_BLQ + U_Q1*z + U_Q2*(z*z) + beff, s2e);
    random beff ~ normal(0,Block_Var) subject=block out=blups;
    predict L_BLQ + U_Q1*z + U_Q2*(z*z) out=ppp;
run;
```

Figure 4. A) Quadratic polynomial mixed model with heterogeneous variance. B) Broken-line linear mixed model with heterogeneous variance. C) Broken-line quadratic mixed model with heterogeneous variance.

of the dose–response relationship were not fully nested among themselves. Therefore, a likelihood ratio test–based approach to model comparison would not have been appropriate. Furthermore, when comparing fit between models fitted using GLIMMIX and NLMIXED procedures, it is important to pay special attention to default software specifications, thereby ensuring that the underlying methods of estimation are aligned so as to enable meaningful comparisons of information criteria. The default specification for method of estimation in GLIMMIX is REML, whereas in NLMIXED, it is ML; therefore, one should explicitly specify ML-based inference in GLIMMIX by indicating method = MSPL (Fig. 4A) in SAS software code.

RESULTS AND DISCUSSION

Base Mixed Model: Implementation and Inference

After fitting a base mixed model to the response G:F assuming a homogeneous residual variance (as in Eq. [1]), we assessed model assumptions using a plot of studentized residuals over levels of SID Trp:Lys ratios (Fig. 3A). The residual plot indicated no evidence

for extreme observations beyond probabilistic expectation, as all studentized residuals were within ±2.5. Yet studentized residuals seemed to be more dispersed around their expected mean 0 for some SID Trp:Lys ratio treatments (i.e., 14.5 and 24.5%) than for others (i.e., 16.5 and 18.0%) thereby questioning the standard assumption of a homogeneous residual variance across all treatments. Instead, the plot of studentized residuals (Fig. 3A) suggested that the residual variance might differ among treatments. To address this departure from model assumptions, we expanded our base mixed model to explicitly accommodate heterogeneous residual variances. Table 1 shows BIC statistics for model fit assessment for alternative base mixed models fitted either with a homogeneous residual variance (i.e., a common variance across SID Trp:Lys ratios) or with heterogeneous residual variances. Heteroskedastic groups were defined either on each SID Trp:Lys ratio (i.e., 7 variance groups) or on 2, 3, or 4 so-called "variance groups" consisting of empirical combinations of SID Trp:Lys ratio treatments. Base mixed models with heterogeneous residual variances for 2, 3, or 4 groups of treatments fitted the data better than the base mixed model with a common homogeneous residual variance;
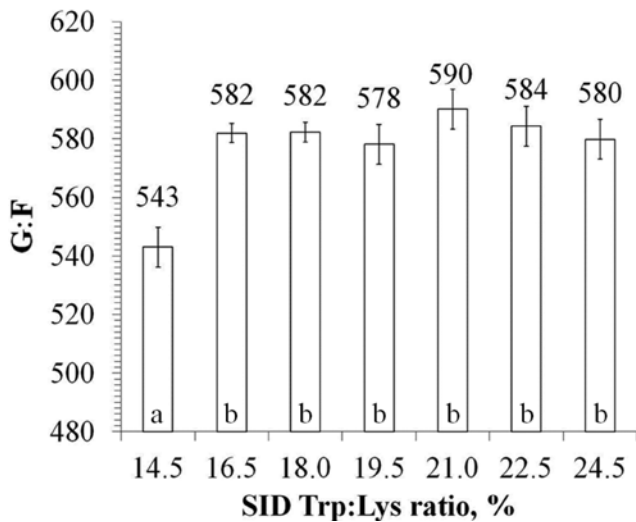
Figure 5. Estimated least squares means of G:F (SEM), expressed in grams per kilogram, for experimental diets consisting of increasing levels of standardized ileal digestible (SID) Trp:Lys ratio using the selected base mixed model with heterogeneous residual variances. [a,b]Different letters indicate significant mean differences between treatments ($P < 0.05$).

this assessment was supported by smaller values of the BIC statistic (Table 1). Therefore, we selected for further modeling steps the most parsimonious model, that is, the base mixed model with fewest variance components that best fit the data, in this case the model with heterogeneous residual variances for 2 groups of SID Trp:Lys ratios consisting of 16.5 and 18.0% vs. 14.5, 19.5, 21.0, 22.5, and 24.5%. Indeed, a plot of studentized residuals obtained from fitting the selected base model with heterogeneous residual variances suggested a more even spread of residuals for all SID Trp:Lys ratio treatments (Fig. 3B). Figure 5 shows the estimated least squares means of G:F for experimental diets consisting of increasing levels of SID Trp:Lys ratio allowing for heterogeneous residual variance across 2 groups, as specified based on the best fitting base mixed model.

Dose–Response Models

As previously indicated, our implementation of nonlinear dose–response mixed models included elicitation of initial parameter values to facilitate the estimation process. Using parameter estimates from the base mixed model (Fig. 5), we specified initial values for the scaled plateau level (i.e., φ_{BLL} or φ_{BLQ}) at approximately 582 g BW gain/kg feed intake whereas the slope β_{BLL} for the linear segment of the BLL model was approximated at 1,950 (calculated using values from Fig. 5 as $(582 - 543)/(0.165 - 0.145)$). It is noted that the sign of the initial value provided for the slope β_{BLL} must be multiplied by -1 to yield the value $-1,950$ given the model parameterization process implemented in SAS software. For breakpoint

parameters, namely ω_{BLL} or ω_{BLQ}, initial values were specified at 15.0, 16.0, and 17.0% SID Trp:Lys ratios based on descriptive assessments of the data. Initial values for variance components were obtained from corresponding estimates of variance components from the fitted base mixed model and specified at 11 for BW block (σ_b^2), 56 for variance group 1 ($\sigma_{e_1}^2$), and 268 for variance group 2 ($\sigma_{e_2}^2$).

To facilitate model convergence, we set up a grid search across initial parameter values. For a grid search, each model parameter was assigned an array of initial values selected within the domain of each parameter; this grid search approach can be particularly important for variance components. For the BLL model, we set an array of parameter values consisting of all combinations of $\varphi_{BLL} = [578, 582, 586]$, $\beta_{BLL} = [-975, -1,950, -3,900]$, $\omega_{BLL} = [0.15, 0.16, 0.17]$, $\sigma_b^2 = [6, 11, 22]$, $\sigma_{e_1}^2 = [28, 56, 112]$, and $\sigma_{e_2}^2 = [134, 268, 536]$. Eliciting an array of reasonable initial values for each parameter can be challenging. One possible approach may be to halve or double the individual values specified to initiate the iterative estimation process, provided that all values in the grid fall within the bounds of the parameter space. Following a similar rationale, we specified initial values for parameters of the BLQ model. The initial values for $\beta_{2,BLQ}$ were informed based on the β_2 coefficient from the QP model, whereby $\beta_{2,BLQ} = [-4,685, -9,369, -18,738]$.

After fitting the BLL and BLQ models, convergence was reached and all parameter estimates were found to be within the plausible range specified by the grid search and away from the extreme values of each grid search. It is noted that the array of initial values that initiates the grid search can be lengthened and/or tuned using a trial-and-error approach to ensure that, ultimately, the point estimate of the parameter of interest falls within the specified array grid. Such tuning is important to minimize the chances of convergence to a local maxima. As such, more than 3 initial values may be needed for some parameters to enhance the search through the likelihood surface. Yet one should keep in mind that larger grids can substantially increase the use of computational resources. Furthermore, it is worth noticing that, even with all these precautions, convergence to suboptimal local maxima can still occur. Therefore, there is no substitute for a careful critical assessment of results in the specific context of a given data application.

After model fitting, the heteroskedastic versions of the QP, BLL, and BLQ models yielded BIC of 353.7, 343.4, and 345.2, respectively, indicating that the BLL model was the better fitting one. For further characterization, we present fitted equations and interpretations based on all competing dose–response mixed models.

We note, however, that for the purpose of reporting results in scientific publications, it is often recommended that only the best fitting model or, alternatively, models with comparably better fit (i.e., lowest BIC values with differences smaller than 2 points) be used to draw conclusions, to avoid misleading readers using inference from models with relatively poorer fit to the data.

Figure 6 shows fitted regression lines based on all 3 competing dose–response mixed models overlaying a scatter plot of the data. The estimated regression equations for these models are presented below followed by the corresponding estimated SEM for each parameter estimate in the succeeding line:

Quadratic polynomial predictive equation:

$$G{:}F = 192.67 + 3{,}859.90 \times (\text{Trp:Lys ratio}) - 9{,}369.19 \times (\text{Trp:Lys ratio})^2$$

(SEM 94.11, SEM 974.12, and SEM 2,477.57, respectively).

Broken-line linear predictive equation:

$$G{:}F = 582.68 - 1{,}950.00 \times (0.1653 - \text{Trp:Lys ratio}),$$
if SID Trp:Lys ratio < 16.5%
(SEM 2.41, SEM 347.24, and SEM 0.0018, respectively), and
$$G{:}F = 582.68$$
if SID Trp:Lys ratio $\geq$ 16.5%
(SEM 2.41).

Broken-line quadratic predictive equation:

$$G{:}F = 582.43 - 1{,}950.00 \times (0.1604 - \text{Trp:Lys ratio}) - 40{,}000.00 \times (0.1604 - \text{Trp:Lys ratio})^2$$

(SEM 2.05, SEM 0.00019, SEM 0.0021, SEM 0.000002812, and SEM 0.0021, respectively), if SID Trp:Lys ratio < 16.0%, and
$$G{:}F = 582.43$$
if SID Trp:Lys ratio $\geq$ 16.0%
(SEM 2.05).

Based on its large BIC value, we concluded that the QP model showed the poorest data fit of all. This is apparent in Fig. 6, where the estimated regression line for the QP model appeared to underestimate G:F at the 16.5% Trp:Lys ratio while seemingly overestimating it at the 14.5 and 19.5% Trp:Lys ratios. Furthermore, the functional form imposed by the QP model on these data forced a maximum predicted G:F at approximately the 21% Trp:Lys ratio followed by a decrease in predicted G:F above this nutrient level. Such predicted decline in G:F at the highest levels of Trp:Lys ratios is not consistent with the data considered in this study (Fig. 6). To further support this point, it is noted that there were no significant differences in mean G:F between treatments consisting of a 21.0, 22.5, or 24.5% Trp:Lys ratio ($P > 0.90$ in all cases; Fig. 5).
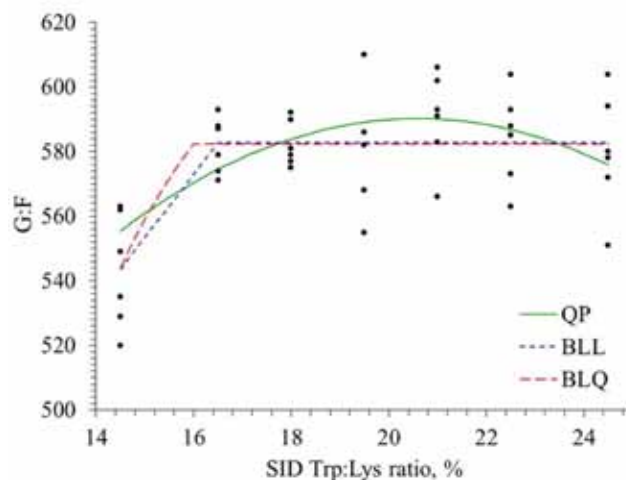

Figure 6. Fitted regression lines for G:F, expressed in grams per kilogram, as a function of experimental diets consisting of increasing levels of standardized ileal digestible (SID) Trp:Lys ratio for competing dose–response linear and nonlinear mixed models accounting for heterogeneous residual variances, including a quadratic polynomial (QP) model, a broken-line linear (BLL) ascending model, and a broken-line quadratic (BLQ) ascending model.

The best fitting model for this data set, that is the BLL model, estimated the SID Trp:Lys ratio breakpoint at 16.53% (95% CI [16.07, 17.00]). Figure 7 illustrates the fitted BLL model, including the breakpoint estimate and its corresponding estimated 95% CI. A note of caution is in order here: Although the BLL breakpoint parameter was technically identifiable in this case, it is apparent from this data example that a very limited amount of information was available to infer on this parameter. In other words, one need only 2 anchoring points to uniquely identify a straight line. In this data, 2 anchoring points were available to trace the ascending linear section of this piecewise BLL regression, namely the 14.5 and 16.5% SID Trp:Lys ratio, both below the estimated BLL breakpoint at 16.53%. So, although technically valid, the BLL breakpoint estimate is noted to be based on the bare minimum amount of information needed to draw a straight line; this should raise questions about the stability of the BLL breakpoint estimate. Clearly, an experiment will have to be designed in which numerous levels of SID Trp:Lys ratio varying between 14.5 and 16.5% are fed, to better characterize the functional form of the G:F trend across this range of SID Trp:Lys ratios.

The BLQ model initially appeared to have an intermediate fit to the data relative to the BLL model and the QP model; in fact, its BIC value seemed to support the BLQ model as a closely ranked second to the best-fitting BLL model. However, on closer evaluation, it should look suspicious that the BLQ breakpoint estimate (i.e., 16.04%) fell below the 16.50% Trp:Lys ratio mark. That is, the quadratic piece of the fitted BLQ model rested on just 1 anchoring point along the *x*-axis, namely the 14.50% Trp:Lys ratio (Fig. 6). In other
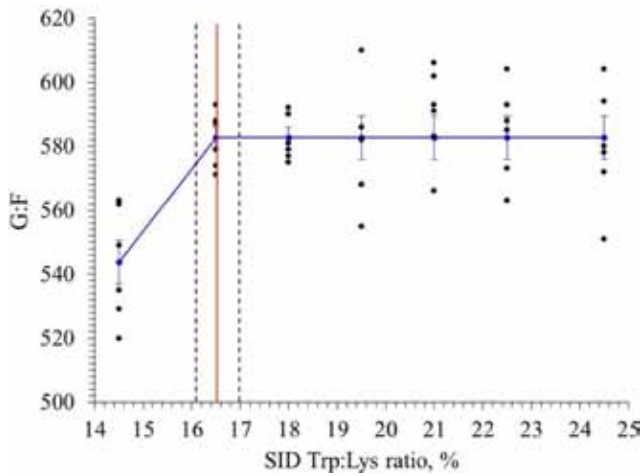
Figure 7. Nonlinear broken-line linear ascending mixed model for G:F, expressed in grams per kilogram, accounting for heterogeneous residual variances, including mean predictions (ascending and horizontal lines), 95% confidence interval on the mean (whiskers), and estimated standardized ileal digestible (SID) Trp:Lys ratio breakpoint (vertical line at 16.5% SID Trp:Lys ratio) with corresponding 95% confidence interval (vertical dashed lines; [16.1, 17.0]).

words, the data contained no information to assess curvature of a fitted line for G:F between the 14.50 and 16.50% Trp:Lys ratios. Indeed, BLQ models fitted with breakpoint parameters held constant in this range of X have unchanged −2 log likelihoods (data not shown). As a result, the BLQ breakpoint estimate obtained here should not be considered informative or meaningful but rather suspect and questionable. In fact, the numerical value obtained as a BLQ breakpoint estimate may be interpreted as a reflection of problems in the estimation process. This is further supported by warning messages obtained in the SAS log when fitting the BLQ model. Even more dramatically than with the BLL model, the end conclusion here is the need for a follow-up study designed to feed multiple levels of SID Trp:Lys ratio in the range between 14.50 and 16.50%.

It is important to note that dose–response models should be fitted and plotted at the level of the experimental unit so that the underlying variability in the data is properly accounted for. In other studies, dose–response models were fitted to the average values of each treatment for a given response variable (Robbins et al., 2006). This practice may appear to enhance model fit, although at the price of disregarding the underlying experimental error. As a result, inferential precision would be artificially exaggerated (i.e., deflated estimates of uncertainty) with subsequent undesirable consequences on Type I error and repeatability of results.

Mixed Model Selection

Commonly used fit statistics in the context of mixed models include the BIC and the Akaike's information

criterion (**AIC**), among others (Milliken and Johnson, 2009). For these information criteria, smaller values are indicative of models with better fit to the data. In this study, we used a ML-based BIC as our fit statistic of choice to select between competing dose–response mixed models. The calculation of BIC yields fit criteria that are slightly more conservative than that of the AIC, as the BIC tends to put a greater penalty on number of model parameters, thus favoring more parsimonious models (Schwarz, 1978). For any given model, the value of the BIC can take any number on the real line from minus infinite to positive infinite. Yet the BIC value is only meaningful in relative terms, that is, when comparing 2 or more models applied to the same data set (Milliken and Johnson, 2009). A model with a smaller BIC value indicates better fit to the data; more specifically, models that differ in their BIC values by at least 2 points are considered to have meaningful differences in their data fit (Raftery, 1996). Only the best fitting model, that is, the model with the smallest BIC, should be used for estimation and inference, as models with poorer fit could yield artifactual, even apparently contradicting, conclusions.

Another model selection criterion commonly used to assess goodness of fit in the context of fixed-effects regression models is the coefficient of determination (R^2). As intuitive an interpretation of goodness of fit as it might have (i.e., proportion of variability explained by an effect), R^2 is often misused in the context of mixed models (Littell et al., 2006). More specifically, for models that have more than one variance component (i.e., mixed models), one should recall that R^2 is not uniquely defined (Kvålseth, 1985) and that the variety of alternative specifications of R^2 is not equivalent (Kvålseth, 1985). This, in turn, can lead to nontrivial pitfalls during interpretation of data analysis results (Kvålseth, 1985; Willett and Singer, 1988).

Homogeneous vs. Heterogeneous Residual Variances

In the context of linear models, it is well described that violation of the homogeneous residual variance assumption poses a considerable risk to inference, probably more so than violations of the normality assumption on residuals (Milliken and Johnson, 2009). Incorrectly assuming homogeneous residual variances due to unchecked residual assumptions can impair inferential efficiency when treatment means (or functions thereof) are of interest (Wiggans and VanRaden, 1991; Wolfinger, 1996). In this study, when homogeneous residual variances were incorrectly assumed, the breadth of the BLL-based 95% CI on the estimated SID Trp:Lys ratio breakpoint for maximum G:F was increased by approximately 44% relative to that of the BLL model with heterogeneous residual variances

(95% CI [15.8, 17.4] vs. [16.1, 17.0], respectively). This example of inappropriately calibrated inferential uncertainty illustrates typical inferential inefficiencies associated with erroneous assumptions in the context of dose–response relationship mixed models.

Hierarchy Principle in Polynomial Models

On building of polynomial regression models, abiding by the hierarchy principle is generally recommended to ensure proper model formulation (Peixoto, 1987). The hierarchy principle states that if a higher-order polynomial term is retained in the model, then the related lower-order terms should also be kept in the model whether or not the coefficients for these lower-order terms are significant (Kutner et al., 2005). This recommendation has not always been heeded in the animal sciences, in particular in the context of estimation of nutritional requirements. For example, BLQ models lacking a first-order linear term have been used (Robbins et al., 2006; Pesti et al., 2009) counter to the hierarchy model building principle. Only hierarchically well-formulated models are invariant under linear transformation; otherwise, significance tests on regression coefficients can yield artifactual results (Peixoto, 1987). Artifactual results may not be obvious for any given data applications; therefore, the importance of taking this principle into careful consideration during model building.

Troubleshooting Nonlinear Mixed Models

Several useful tips on troubleshooting implementation of nonlinear models can be found in the literature (Kiernan et al., 2012). Specific to nonlinear mixed models such as BLL and BLQ, some of the most common issues include convergence of the iterative estimation process and estimation failures.

Use of a Grid Search to Facilitate Estimation. One of the biggest challenges in nonlinear mixed models is to reach convergence of the iterative estimation process. Even when convergence is attained, there is no guarantee that it be on a global maxima (optimal solution) of the likelihood function as opposed to a local maxima (suboptimal solution). Most nonlinear optimization methods can be trapped by local maxima and provide no clear indication of convergence to suboptimal solutions. The use of a grid search approach can assist in this process by providing plausible initial parameter values over the range of the feasible parameter space, thus more efficiently guiding the iterative parameter estimation process toward optimal global solutions and away from suboptimal ones (Kiernan et al., 2012). Specification of initial values is particularly important for variance components as well as for nonlinear parameters, such

as the breakpoint parameter in the piecewise regression. Initial values can be specified in NLMIXED by using the PARMS statement (Fig. 4B and 4C). The actual initial values inputted into NLMIXED can be adjusted based on preliminary analysis and also to assess sensitivity of the final inference to starting values. For the data sets that we have used, we find that specification of at least 3 initial values for each parameter facilitates a grid search over the likelihood surface. The START option in the NLMIXED procedure will call for an output, which shows exactly which initial value the procedure chose for each of the parameters. A greater number of initial values in the grid search will enhance the search through the likelihood surface, particularly for parameters of primary interest (i.e., breakpoints). One should, however, also consider that large grid searches can substantially increase the use of computational resources.

Other Troubleshooting Techniques. Besides using a grid search, failures to converge can be minimized by 1) specifying reasonable models that reflect the functional form of the dose–response appropriately for a given data set, whereby "reasonable models" can be informed from preliminary analysis or fitting a base model that does not assume any functional form in the relationship between response and treatment, as in our case; 2) increasing the number of maximum iterations before which the estimation process will abort if convergence has not been reached (i.e., in the NLMIXED procedure, this can be specified with the option MAXITER); 3) defining reasonable boundaries for selected parameters, in particular the breakpoint parameter, which should not lie outside of the range of dose levels considered in a given study (i.e., in the LNMIXED procedure, this can be specified with the BOUNDS statement); 4) careful tuning of the convergence criterion (i.e., in NLMIXED, this can be attained by modifying the relative gradient criterion using the option GCONV; in particular, GCONV = 0 will force the procedure to continue to a greater number of iterations until the relative change in a succeeding convergence criterion, namely the function value criterion (FCONV) is met), as suggested by Kiernan et al. (2012); and 5) evaluating the programing code for bugs and coding errors, as with implementation of any statistical software.

Last but not least, we emphasize that neither specifying grid searches nor any of the troubleshooting tips discussed this far are infallible, as it was apparent for the BLQ model fitted in this study. In fact, none of these troubleshooting techniques can by themselves guarantee proper use and behavior of the estimation process. In other words, fitting of nonlinear mixed models should not be approached as a plug-and-play software exercise. There is no substitute for a careful critical assessment of estimation and inference results in the particular context of a problem of interest.

Summary

This paper presents concepts underlying the implementation of linear and nonlinear mixed models for dose–response relationships accounting for correlated data structure and heterogeneous residual variances. We illustrated the inferential implications of properly checking and addressing model assumptions, particularly pertaining to the assumption of a common residual variance, as well as the use of the hierarchy principle for model building. Additionally, we explained the importance of using proper fit statistics for model selection in the context of mixed models and the pitfalls of using the conventional fixed-effect-based coefficient of determination (R^2). We further demonstrated practical approaches to facilitate some of the computational challenges associated with fitting of nonlinear mixed models, including use of a grid search to facilitate convergence of the iterative estimation process.

LITERATURE CITED

Cernicchiaro, N., D. G. Renter, S. Xiang, B. J. White, and N. M. Bello. 2013. Hierarchical Bayesian modeling of heterogeneous variances in average daily weight gain of commercial feedlot cattle. J. Anim. Sci. 91:2910–2919. doi:10.2527/jas.2012-5543

Gonçalves, M. A. D., S. Nitikanchana, M. D. Tokach, S. S. Dritz, N. M. Bello, R. D. Goodband, K. J. Touchette, J. Usry, J. M. DeRouchey, and J. C. Woodworth. 2015. Effects of standardized ileal digestible tryptophan:lysine ratio on growth performance of nursery pigs. J. Anim. Sci. 93:3909–3918. doi:10.2527/jas.2015-9083

Kiernan, K., J. Tao, and P. Gibbs. 2012. Tips and strategies for mixed modeling with SAS/STAT procedures. In: Proc. SAS Global Forum, Orlando, FL. p. 1–18.

Kutner, M. H., C. Nachtsheim, J. Neter, and W. Li. 2005. Applied linear regression models. 5th ed. McGraw-Hill/Irwin, New York, NY. p. 299–305.

Kvålseth, T. O. 1985. Cautionary note about R^2. Am. Stat. 39:279–285.

Lavagnini, I., and F. Magno. 2007. A statistical overview on univariate calibration, inverse regression, and detection limits: Application to gas chromatography/mass spectrometry technique. Mass Spectrom. Rev. 26:1–18. doi:10.1002/mas.20100

Littell, R. C., G. A. Milliken, W. W. Stroup, R. D. Wolfinger, and O. Schabenberger. 2006. SAS® for mixed models, 2nd ed. SAS Inst. Inc., Cary, NC.

Milliken, G. A., and D. E. Johnson. 2009. Analysis of messy data: Designed experiments. Vol. 1, 2nd ed. CRC Press, Boca Raton, FL.

Peixoto, J. L. 1987. Hierarchical variable selection in polynomial regression models. Am. Stat. 41:311–313.

Pesti, G. M., D. Vedenov, J. A. Cason, and L. Billard. 2009. A comparison of methods to estimate nutritional requirements from experimental data. Br. Poult. Sci. 50:16–32. doi:10.1080/00071660802530639

Raftery, A. E. 1996. Approximate Bayes factors and accounting for model uncertainty in generalized linear regression models. Biometrika 83:251–266. doi:10.1093/biomet/83.2.251

Robbins, K. R., A. M. Saxton, and L. L. Southern. 2006. Estimation of nutrient requirements using broken-line regression analysis. J. Anim. Sci. 84:E155–E165.

Robbins, K. R., H. W. Norton, and D. H. Baker. 1979. Estimation of nutrient requirements from growth data. J. Nutr. 109:1710–1714.

Schwarz, G. 1978. Estimating dimension of a model. Ann. Stat. 6:461–464. doi:10.1214/aos/1176344136

Vedenov, D., and G. M. Pesti. 2007. A comparison of methods of fitting several models to nutritional response data. J. Anim. Sci. 86:500–507. doi:10.2527/jas.2007-0536

Wiggans, G. R., and P. M. VanRaden. 1991. Method and effect of adjustment for heterogeneous variance. J. Dairy Sci. 74:4350–4357. doi:10.3168/jds.S0022-0302(91)78631-1

Willett, J. B., and J. D. Singer. 1988. Another cautionary note about R^2: Its use in weighted least-squares regression analysis. Am. Stat. 42:236–238.

Wolfinger, R. D. 1996. Heterogeneous variance covariance structure for repeated measures. J. Agric. Biol. Environ. Stat. 1:205–230. doi:10.2307/1400366