

Techinal Note: Estimating Parameters of Nonlinear Segmented Models

J. G. Fadel

Department of Animal Science
University of California, Davis 95616

ABSTRACT

The objective of this technical note is to develop an applied technique to estimate parameters using the Statistical Analysis System's nonlinear procedure (SAS PROC NLIN) for segmented models that have a change point or lag as one of the parameters. The goal is to select good starting values for the parameters near the global minimum or at least near a lower local minimum compared with what might be achieved using traditional starting values. The model used was $f_1(t) = B1 + B4$ if $t \leq B3$ and $f_2(t) = B1 \cdot \exp[-B2 \cdot (t - B3)] + B4$ if $t > B3$, where $B1$, $B2$, $B3$, and $B4$ are the parameters that require estimates, $B3$ is the change point or lag, and t = time. This technical note illustrates the solution when a traditional grid search for starting values is used and demonstrates a modified technique where starting values are systemically determined by fixing $B3$ over a range of reasonable values and then using the parameters from the solution with the lowest residual sums of squares as the starting values for the final solution. The modified method resulted in a lower square root of mean square error compared with the traditional method. The estimates for $B3$ (lag) were 3.5 for the modified method compared with 4.5 for the traditional method. This technique works well when using the SAS PROC NLIN procedure but can be modified to work with other statistical packages.

(Key words: change point, lag, nonlinear, segmented)

Nonlinear models with a change point or lag are also called segmented models and are often used to describe fiber digestion as well as other biological processes. An example of this type of model is:

$$E[F_1(t)] = f_1(t) = B1 + B4 \text{ if } t \leq B3 \quad [1]$$

$$E[F_2(t)] = f_2(t) = B1 \cdot \exp[-B2 \cdot (t - B3)] + B4 \text{ if } t > B3 \quad [2]$$

A model of this type could be used to describe fiber remaining at time t of fermentation, where $B1$ = available fiber for digestion, $B2$ = rate of fiber digestion (t^{-1}), $B3$ = lag time or change point (t), and $B4$ = the indigestible fraction. The data can be described by n points ($P_1 \dots P_n$) and may be unequally spaced. The two coordinates of the i th point are described by $P_i = (t_i, d_i)$, where t_i is the measured time point and d_i is the measured value or fiber remaining at the i th time.

The problem is how to partition the observation points between equation [1] and [2], for a given data set, to estimate $B1$, $B2$, $B3$, and $B4$ such that

$$\sum_{i=1}^k [d_i - f_1(t_i)]^2 + \sum_{i=k+1}^n [d_i - f_2(t_i)]^2 \quad [3]$$

is minimized, where $t_k \leq B3 < t_{k+1}$.

The segmented model (equation [1] and [2]) is continuous but not smooth. Continuity implies equation [1] and [2] have the same solution when t equals $B3$ ($f_1[B3] = B1 + B4$ and $f_2[B3] = B1 + B4$). Smoothness implies the first derivatives of equation [1] and [2] with respect to t are identical when t equals $B3$ ($df_1[t]/dt = 0$, and $df_2[t]/dt = -B2 \cdot B1 \cdot \exp[-B2 \cdot (t - B3)] = -B2 \cdot B1$). The parameters can be estimated with statistical packages that have a nonlinear estimation procedure, such as SAS (SAS Inst., Inc., Cary, NC). The procedure used in SAS is PROC NLIN, and this procedure estimates parameters from segmented models that are continuous and smooth. Although the segmented model described by equation [1] and [2] does not meet the criteria for segmented models in SAS, this procedure is widely used. Minimization algorithms are derived under the assumption that the first partial derivative of residual sums of squares, with respect to the parameters, is smooth and continuous.

Selection of good starting values is important to obtain good nonlinear regression solutions. According to Motulsky and Ransnas (1987), "A poor selection of initial values may have several consequences: 1) in a well-behaved system, the amount of computer time required to reach a solution will be increased, but the solution will be correct; 2) in other situations, poorly selected initial values can lead the nonlinear regression program to go in the wrong direction or never converge on

Received December 19, 2002.

Accepted June 9, 2003.

E-mail address: jgfadel@ucdavis.edu.

a solution; 3) it is also possible that poorly selected initial values can cause the program to converge on the wrong solution, a local minimum. The choice of initial values is less important with equations containing only one or a few parameters than with equations containing many parameters." If case 3 occurs, the experimenter can usually suspect a wrong solution because the coefficient of determination is low or the standard errors of the estimates are high, but these evaluations do not always uncover wrong solutions. Problems of identifiability and ill-conditioning have been summarized by Seber and Wild (1989). Initial starting values generally are very important, and especially so with segmented models. Good overviews of nonlinear regression can be found in Bates and Watts (1988) and Seber and Wild (1989). An applied SAS technique is not available that allows the selection of good starting values in a systematic, reliable procedure. However, transition functions have also been used to solve the change point problem (Boston et al., 2000).

The objective of this article is to illustrate a technique to find good starting values around a local minimum that is lower or equal to the local minimum found using traditional methods. This technique will be illustrated by using a specific nonlinear segmented model and should work with other nonlinear segmented models.

Much has been written about change point problems (Carlstein et al., 1994; Sinha et al., 1995), but solutions to applied problems using statistical packages are limited. One solution to this problem is to fix B3 (lag) across a range of reasonable B3 values and solve for B1, B2, and B4. Then, the solution set with the lowest residual sums of squares is selected and the fixed B3 and the estimates of B1, B2, and B4 are used as starting values for the final solution. The selection of the reasonable range should be dictated somewhat by the experimenter's knowledge of the system. The number of fixed B3 to use was selected to be every 0.1 units from 0 to 8. However, an incremental unit of 0.1 was selected arbitrarily and the exact increment to use is more of a mathematical problem and is outside the range of this technical note. The advantage of this modified approach is that one should achieve a better solution than by selecting reasonable starting values for all parameters. The datum used for this technical note is a subset of data from Fadel (1992) and is given in Table 1. A traditional SAS program with a reasonable starting grid is under "Parameters" in Appendix 1. A modified SAS program to estimate the parameters is in Appendix 2. The modified SAS program generates a series of datasets using a different fixed lag or B3 for each dataset, and then B1, B2, and B4 are estimated for each fixed B3. The parameters estimated from the solution with

Table 1. Datum for SAS nonlinear regression.

Time	Observation ¹
0	0.5966
0	0.5991
0	0.5918
1	0.5978
1	0.5969
1	0.5937
2	0.5910
2	0.6016
2	0.5902
4	0.5914
4	0.5761
4	0.5768
8	0.5366
8	0.5172
8	0.5450
16	0.3951
16	0.4002
16	0.4028
32	0.3144
32	0.3221
32	0.3171
64	0.3039
64	0.2834
64	0.2731
128	0.2212
128	0.2178
128	0.2301

¹Neutral detergent fiber remaining at time *t*.

the lowest residual sums of squares and the fixed B3 are used as the starting values for the final analysis.

Figure 1 was generated from the first PROC NLIN in Appendix 2 by using selected fixed B3 values near the two minima. Figure 1 shows one local minimum at about B3 = 3.5, and another local minimum at about B3 = 4.5. It is hoped that the local minimum at B3 = 3.5 is the global minimum, but the modified technique cannot guarantee the residual sums of squares is a global minimum. The parameters estimated (B1, B2,

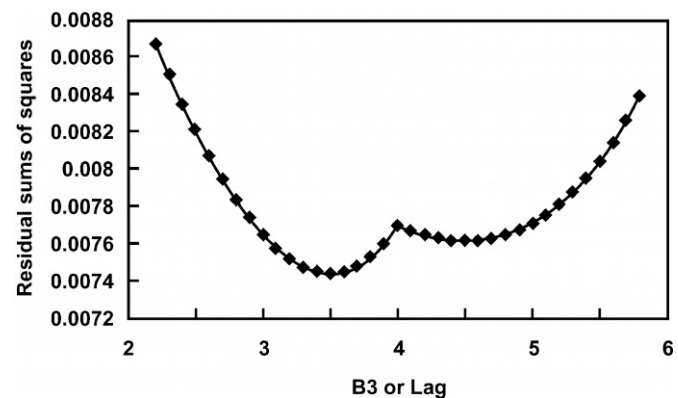


Figure 1. Two minima from several analyses. The parameters B1, B2, and B4 are estimated for each fixed B3 using $f_1(t) = B1 + B4$ if $t \leq B3$ and $f_2(t) = B1 \cdot \exp[-B2 \cdot (t - B3)] + B4$ if $t > B3$.

Table 2. Results from traditional and modified SAS NLIN analyses of datum from Table 1.¹

	B1	B2	B3	B4	RMSE ²
Traditional					
Estimates	0.3425	0.0624	4.5034	0.2495	0.01820
Asymptotic SE	0.00988	0.00742	0.86040	0.00837	
Asymptotic correlations					
B1		-0.5244	-0.3768	-0.8469	
B2			0.7149	0.6192	
B3				0.2654	
Modified					
Estimates	0.3493	0.0564	3.4963	0.2461	0.01798
Asymptotic SE	0.01030	0.00515	0.57710	0.00833	
Asymptotic correlations					
B1		-0.5223	-0.4343	-0.8116	
B2			0.4773	0.6436	
B3				0.1559	

¹Traditional indicates that the analysis used starting values from a predetermined grid. "Modified" indicates analysis used starting values based on lowest residual sums of squares when B1, B2, and B4 were estimated at varying B3 or lag values. Equations used to estimate B1, B2, B3, and B4: $f_1(t) = B1 + B4$ if $t \leq B3$ and $f_2(t) = B1 * \exp[-B2 * (t - B3)] + B4$ if $t > B3$.

²RMSE = square root of mean square error.

and B4) are assumed to approach their optimized values at each fixed B3. However, the parameter estimates might not be optimum, and the experimenter needs to be aware of other techniques to test their models, as described elsewhere (Bates and Watts, 1988; Seber and Wild, 1989). Table 2 contains information based on the results from Appendix 1 and from the second PROC NLIN in Appendix 2. The results in Table 2 show the square root of mean square error for the traditional and modified techniques are 0.01820 and 0.01798, respectively. The estimates of B2 (rate of digestion) and B3 (lag) using the traditional analysis may lead to erroneous interpretation of the datum. The estimates of B2 and B3 are 0.062 and 4.5 for the traditional analysis compared with 0.056 and 3.5 for the modified analysis. Not all datasets will have more than one local minimum, but the technique described in this paper provides another method in an attempt to reach a lower local minimum than that found by using a traditional approach in selecting starting values. The frequency of other datasets that would benefit by a lower root mean square error from the application of this modified technique is not known, and a careful analysis to provide relevant statistics is outside the scope of this technical note. However, 20 datasets were generated from the dataset in Table 1 using various standard deviations, and the modified procedure worked in every case and provided lower or equal residual sums of squares compared with the traditional procedure. The experimenter should consider using the technique illustrated in this manuscript if the parameter estimates do not seem to make biological sense or do not compare with other replicates. Additionally, this modified technique should be used if one is interested in estimates of the rate

constant (B2) and the lag (B3). If one is interested in the mean retention time primarily, then the traditional and modified techniques may provide similar results. For example, in this article, the percentage change from the traditional to the modified technique is about 10% for B2 and 22% for B3. However, the percentage change is only 3.4% for the mean retention time ($1/B2 + B3$). Although the technique developed in this article applies to SAS PROC NLIN, it can be modified to work with other statistical packages.

ACKNOWLEDGMENTS

The author wishes to thank L. Gustafsson for fruitful discussions concerning solutions to change point problems.

REFERENCES

- Bates, D. M., and D. G. Watts. 1988. *Nonlinear Regression Analysis and Its Applications*. John Wiley & Sons, Inc., New York, NY.
- Boston, R. C., D. G. Fox, C. Sniffen, E. Janczewski, R. Munson, and W. Chalupa. 2000. The conversion of a scientific model describing dairy cow nutrition and production to an industry tool: The CPM dairy project. Pages 361–377 in *Modelling Nutrient Utilization in Farm Animals*. J. P. McNamara, J. France, and D. Beever, eds. CABI Publishing, Oxon, U.K.
- Carlstein, E., H. G. Müller, and D. Siegmund, eds. 1994. *Change-point Problems*. IMS Lecture Notes-Monograph Series, Institute of Mathematical Statistics, Hayward, CA.
- Fadel, J. G. 1992. Application of the theoretical optimal sampling schedule designs for fiber digestion parameter estimation in sacco. *J. Dairy Sci.* 75:2184–2189.
- Motulsky, H. J., and L. A. Ransnas. 1987. Fitting curves to data using nonlinear regression: A practical and nonmathematical view. *FASEB J.* 1:365–374.
- Sinha, B., A. Rukhin, and M. Ahsanullah, ed. 1995. *Applied Change Point Problems in Statistics*. Nova Science Publishers, Inc., Commack, NY.

Seber, G. A., and C. J. Wild. 1989. Nonlinear Regression. John Wiley & Sons, New York, NY.

APPENDIX 1

Traditional SAS program to determine estimates for nonlinear segmented model.

```
LIBNAME mylib "";
*Read in datum;
DATA mylib.data;
INPUT time NDF;
CARDS;
*Datum from Table 1 goes here;
;
PROC NLIN DATA=mylib.data BEST=100 MAXITER=50 G4SINGULAR
  METHOD=DUD OUTEST=EST ; OUTPUT OUT=mylib.results1 P=predy R=resid;
TITLE 'JDSAppendix 1';
PARAMETERS
B1 = .2 .3 .4
B2 = .02 .05 .1
B3 = .1 5 8.
B4 = .2 .25 .3;
IF time >= B3 THEN DO;
TT = time - B3;
E = EXP(-B2*TT);
MODEL NDF = B1*E + B4;
END;
IF time < B3 THEN DO;
MODEL NDF = B1 + B4;
END;
*Use following to save residual sums of squares or _SSE_;
DATA mylib.results2 (KEEP= _SSE_ B1 B2 B3 B4);
SET EST;
IF _TYPE_ = "FINAL" then OUTPUT mylib.results2;
*Use proc summary to save corrected total sums of squares;
PROC SUMMARY DATA=mylib.results1; VAR NDF; OUTPUT OUT = mylib.results3 CSS=ctss;
*Merge data sets;
DATA finalresults;
MERGE mylib.results2 mylib.results3;
PROC PRINT;
RUN;
```

APPENDIX 2

Modified SAS program to determine estimates for nonlinear segmented model.

```
LIBNAME mylib "";
*Read in datum;
DATA mylib.data1; INPUT time NDF;
CARDS;
*Datum from Table 1 goes here;
;
*Create new dataset with Lag (B3) as fixed;
DATA mylib.data2; SET mylib.data1; DO B3 = 0 to 8 by .1; OUTPUT; END;
PROC SORT; by B3;
*Run first Proc NLIN to find lowest _SSE_ to find starting lag (B3) value for final solution;
*Fix lag (B3) and just estimate 3 parameters (B1, B2, and B4);
PROC NLIN DATA=mylib.data2 BEST = 100 MAXITER=50 G4SINGULAR
METHOD=DUD OUTEST=EST ; OUTPUT OUT=mylib.results1 P=predy R=resid; BY B3;
TITLE 'JDSAppendix 2';
PARAMETERS B1 = .3 B2 = .05 B4 = .2;
IF time >= B3 THEN DO;
TT = time - B3; E = EXP(-B2*TT);
MODEL NDF = B1*E +B4;
END;
IF time < B3 THEN DO;
```

```

MODEL NDF = B1 + B4;
END;
DATA mylib.results2 (KEEP=B3 _SSE_ B1 B2 B4); SET EST;
IF _TYPE_ = "FINAL" then OUTPUT mylib.results2;
*Use following plot statement to examine local minimums;
PROC PLOT DATA=mylib.results2; PLOT _SSE_ *B3; PROC print;
*Use following to find solution with lowest residual sums of squares;
PROC SORT DATA=mylib.results2; BY _SSE_;
DATA FINALDATA; SET mylib.results2; RETAIN LOWEST;
LOWEST=MIN(LOWEST,_SSE_);
IF _SSE_ NE LOWEST THEN DELETE; OUTPUT;
PROC PRINT DATA=FINALDATA;
*Set up macros to find starting values for final NLIN;
*Need following DATA statement to save parameters for;
*Use in next PROC NLIN;
*Then pass estimates from previous best fit to starting values in next analysis;
DATA _NULL_; SET FINALDATA;
CALL SYMPUT("BB1",B1);
CALL SYMPUT("BB2",B2);
CALL SYMPUT("BB3",B3);
CALL SYMPUT("BB4",B4);
PROC NLIN DATA=mylib.data1 BEST = 100 MAXITER=50
G4SINGULAR METHOD=DUD OUTEST=EST1;
OUTPUT OUT=mylib.results3 P=predy R=resid;
TITLE 'JDSAppendix 2';
PARAMETERS
B1f = &BB1
B2f = &BB2
B3f = &BB3
B4f = &BB4;
IF time >= B3f THEN DO;
TT = time - B3f;
E = EXP(-B2f*TT);
MODEL NDF = B1f*E +B4f;
END;
IF time < B3f THEN DO;
MODEL NDF = B1f + B4f;
END;
PROC PLOT;
PLOT resid*time="R"/VREF = 0;
PLOT predy*time="P" NDF*time="O"/OVERLAY;
*Use following to save residual sums of squares or _SSE_;
DATA mylib.results4 (KEEP= _SSE_ B1f B2f B3f B4f);
SET EST1;
IF _TYPE_ = "FINAL" then OUTPUT mylib.results4;
*Use proc summary below to save corrected total sums of squares or ctss;
PROC SUMMARY DATA=mylib.results3; VAR NDF; OUTPUT OUT=mylib.results5
CSS=ctss;
*Merge data sets;
DATA finalresults;
MERGE mylib.results4 mylib.results5;
PROC PRINT;
RUN;

```
