

# Is Nonlinear Regression Throwing You a Curve? New Diagnostic and Inference Tools in the NLIN Procedure

Biruk Gebremariam, SAS Institute Inc.

## ABSTRACT

The NLIN procedure fits a wide variety of nonlinear models. However, some models can be so nonlinear that standard statistical methods of inference are not trustworthy. That's when you need the diagnostic and inferential features that were added to PROC NLIN in SAS/STAT<sup>®</sup> 9.3, 12.1, and 13.1. This paper presents these features and explains how to use them. Examples demonstrate how to use parameter profiling and confidence curves to identify the nonlinear characteristics of the model parameters. They also demonstrate how to use the bootstrap method to study the sampling distribution of parameter estimates and to make more accurate statistical inferences.

This paper highlights how measures of nonlinearity help you diagnose models and decide on potential reparameterization. It also highlights how multithreading is used to tame the large number of nonlinear optimizations that are required for these features.

## INTRODUCTION

Nonlinear regression models are formally defined as models in which at least one of the model parameters occurs nonlinearly in the model expression. Consequently, the derivatives of the model expression with respect to the parameters depend on at least one of the parameters. Nonlinear models are used to model complex interrelationships among variables. Common examples of nonlinear models include growth, yield-density, and dose-response models and various models that are used to describe industrial and econometric processes (Ratkowsky 1983; Seber and Wild 1989).

The NLIN procedure fits univariate nonlinear models by using the least squares method. The stochastic structure of models that are fit by the NLIN procedure is given by

$$\begin{aligned} \mathbf{Y} &= \mathbf{f}(\boldsymbol{\theta}; \mathbf{z}) + \boldsymbol{\epsilon} \\ E[\boldsymbol{\epsilon}] &= \mathbf{0} \\ \text{Var}[\boldsymbol{\epsilon}] &= \sigma^2 \mathbf{I}_w \end{aligned}$$

where  $\mathbf{f}$  is the mean function,  $\boldsymbol{\theta}$  refers to the model parameters,  $\mathbf{z}$  is the vector of regressors, and  $\mathbf{Y}$  is the response. The model errors,  $\boldsymbol{\epsilon}$ , are assumed to be uncorrelated. For homoscedastic errors, which are errors with equal yet usually unknown variance,  $\mathbf{I}_w$  is the identity matrix and  $\sigma^2$  is the variance. If the errors are assumed to have unequal variances (heteroscedastic),  $\mathbf{I}_w$  can be a nonidentity diagonal weight matrix.

## BENEFITS OF CLOSE-TO-LINEAR NONLINEAR REGRESSION MODELS

Statistical inference (such as constructing confidence intervals, correlation matrices, and covariance matrices of the parameter estimates) that remains valid under finite sample size is more challenging for nonlinear regression models than for linear regression models.<sup>1</sup> For example, if the error,  $\boldsymbol{\epsilon}$ , is assumed to be independent and identically distributed normal random variables, then the corresponding least squares parameter estimators of a linear regression model are unbiased and normally distributed and have minimum variance among regular estimators (Seber and Wild 1989). These are important properties that standard statistical inferences, which assume normality of the parameter estimates, use in practice. In contrast, least squares parameter estimates of nonlinear regression models acquire these properties only asymptotically as the sample size approaches infinity (Seber and Wild 1989). Consequently, for finite sample size, standard

---

<sup>1</sup>Henceforth, the phrase *nonlinear regression model* stands for the totality of the model expression and the data.

inference might not be trustworthy for a particular nonlinear regression model. In addition, prediction of the response for specified regressor values can suffer from significant bias in nonlinear regression models.

The validity of standard inference in nonlinear regression models depends on both the model expression and the data (Ratkowsky 1983). Close-to-linear nonlinear regression models are those for which standard inference works. In contrast, far-from-linear nonlinear regression models are those for which standard inference does not work. After the parameter estimates are obtained for a particular nonlinear regression model, the first task is to assess whether the nonlinear regression model is a close-to-linear or a far-from-linear model. It is usually considered self-evident that a close-to-linear nonlinear regression model is preferred to a far-from-linear nonlinear regression model (Ratkowsky 1983). The advantages of a close-to-linear model are as follows:

- parameter estimates that are unbiased and have a close-to-minimal variance
- unbiased predicted values of the response
- faster convergence of the least squares minimization
- simplicity of the standard inference in computation, verification, and interpretation

A far-from-linear nonlinear regression model requires additional steps to draw valid statistical inference and prediction. An essential first step is reparameterization of the model,

$$\theta = g(\beta)$$

which consists of defining the old parameters,  $\theta$ , in terms of new parameters,  $\beta$ , by using a one-to-one mapping function,  $g$ . The goal is to obtain a close-to-linear model that has  $\beta$  as its parameters. If the reparameterization yields a close-to-linear model, performing standard inference and prediction with the new model suffices. In nonlinear models that have significant intrinsic curvature, reparameterization might not yield a close-to-linear model (Bates and Watts 1980). Such nonlinear regression models ultimately require the use of inferential approaches such as the bootstrap that, in general, have better finite sample size properties than the standard methods of inference (DiCiccio and Efron 1996)

## DIAGNOSING A FAR-FROM-LINEAR NONLINEAR MODEL IN PROC NLIN

Before SAS/STAT 9.3, the NLIN procedure provided estimates of the confidence intervals, standard errors, and correlation matrix based on the standard inference. It provided no features to check the reliability of these estimates. Neither did it provide alternative options to make more dependable estimates. This is no longer the case. Global measures of curvature, Box's bias, and ODS graphics were added in SAS/STAT 9.3; the PROFILE statement for parameter profiling was added in SAS/STAT 12.1; and the BOOTSTRAP statement for bootstrapping the parameter estimates was added in SAS/STAT 13.1. You can use these options to assess the reliability of the standard inference and to make better statistical inference in cases where the standard inference fails.

In the following sections, a nonlinear regression model based on the Mitscherlich equation,

$$y = \theta_3 + \theta_2 \exp(\theta_1 x)$$

is used to demonstrate the utility of these features of the NLIN procedure. The data set, taken from Clarke (1987), is reproduced using the following statements:

```
data clarke1987;
  input x y;
  datalines;
  1    3.183
  2    3.059
  3    2.871
  4    2.622
  5    2.541
  6    2.184
  7    2.110
```

```

8      2.075
9      2.018
10     1.903
11     1.770
12     1.762
13     1.550
;

```

### Standard Inference in the NLIN Procedure

By default, the NLIN procedure produces statistical results based on the standard inference. The following statements fit a Mitscherlich equation to the preceding data set:

```

proc nlin data=clarke1987;
  parms theta1=-0.15 theta2=2.0 theta3=0.80;
  model y = theta3 + theta2*exp(theta1*x);
run;

```

In addition to showing the parameter estimates, [Figure 1](#) and [Figure 2](#) show the 95% standard confidence intervals and correlation matrix, respectively. The reliability of these estimates is assessed in the following sections. Intuitively,  $\theta_1$ , which occurs nonlinearly in the model expression, is expected to have the least dependable estimates, because the standard inference is based on linear approximation. The next few sections show whether this intuitive expectation holds or not.

**Figure 1** Parameter Estimates and Standard Confidence Intervals  
The NLIN Procedure

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits	
theta1	-0.1031	0.0255	-0.1599	-0.0462
theta2	2.5190	0.2658	1.9268	3.1112
theta3	0.9631	0.3216	0.2466	1.6797

**Figure 2** Standard Correlation Matrix

Approximate Correlation Matrix			
	theta1	theta2	theta3
theta1	1.0000000	0.9233045	-0.9841058
theta2	0.9233045	1.0000000	-0.9723891
theta3	-0.9841058	-0.9723891	1.0000000

### Global Measures of Nonlinearity: The NLINMEASURES Option

Relative intrinsic and parameter-effects curvatures can be used to quantify the global nonlinearity of a nonlinear regression model (Bates and Watts 1980). These measures are scale-invariant with respect to both the parameters and the data. The intrinsic curvature measures the curvature of the expectation surface at the parameter estimates, and the parameter-effects curvature measures the lack of parallelism of parallel parameter lines that are projected onto the expectation surface. As a result, the parameter-effects curvature can be reduced by reparameterization of the model, whereas intrinsic curvature is an inherent property of the model that cannot be affected by reparameterization.

A critical curvature value,  $C_r$ , given by

$$C_r = 1/\sqrt{F(p, n - p; \alpha)}$$

can be used to gauge the statistical significance, at the significance level  $\alpha$ , of the relative intrinsic and parameter-effects curvatures of a nonlinear regression model (Bates and Watts 1980). In the preceding expression,  $n$  and  $p$  stand for the number of observations and number of parameters, respectively. The value of  $C_r$  can be considered the radius of curvature of the  $100(1 - \alpha)\%$  confidence region. A value of either relative curvature that exceeds the critical value signals a far-from-linear model. If the intrinsic curvature exceeds the critical value, it indicates an inherently far-from-linear model that cannot be made a close-to-linear model by reparameterization.

The NLIN procedure provides the NLINMEASURES option in the PROC NLIN statement to produce the nonlinearity measures table. In addition to the maximum intrinsic and parameter-effects curvatures, the nonlinearity measures table contains the root mean square intrinsic curvature, parameter-effects curvature, and critical curvature values. Figure 3 displays this table, which is generated by the following statements:

```
proc nlin data=clarke1987 nlinmeasures;
  parms theta1=-0.15 theta2=2.0 theta3=0.80;
  model y = theta3 + theta2*exp(theta1*x);
run;
```

**Figure 3** Global Measures of Nonlinearity  
The NLIN Procedure

Global Nonlinearity Measures	
Max Intrinsic Curvature	0.0935
RMS Intrinsic Curvature	0.0418
Max Parameter-Effects Curvature	7.0183
RMS Parameter-Effects Curvature	3.1301
Curvature Critical Value	0.5193
Raw Residual Variance	0.0053
Projected Residual Variance	0.0059

Figure 3 shows that for the relative parameter-effects curvature, both the maximum value, 7.0183, and the root mean square value, 3.1301, greatly exceed the critical curvature value, 0.5193. In contrast, the relative intrinsic curvature's maximum value, 0.0935, and the root mean square value, 0.0418, are significantly less than the critical value. These values imply that although the nonlinear regression model is a far-from-linear model, it is not an inherently far-from-linear model because the intrinsic curvature values are well below the critical value. As a result, reparameterizations of the model in terms of new parameters can yield a close-to-linear model.

### Which Parameters Dominate the Global Nonlinearity? The BIAS and HOUGAARD Options

In practice, only a few of the parameters might dominate the global nonlinearity, in which case these parameters are the reparameterization parameters of interest. Unfortunately, the global nonlinearity measures discussed in the preceding section do not differentiate the parameters based on their contribution to the overall curvature. Manifestations of strong nonlinear behavior include significant bias and skewness. Hence, it is essential to estimate at least these basic statistical properties of the parameter estimates in order to identify the reparameterization parameters of interest. You can estimate the bias and skewness of the parameter estimates of a nonlinear regression model by using Box's bias measure and Hougaard's skewness measure (Box 1971; Hougaard 1982). These measures are based on beyond-linear approximation of the expectation surface.

The NLIN procedure provides the BIAS and HOUGAARD options, which add to the parameter estimates table columns that correspond to Box's bias and Hougaard's skewness measures. You can invoke these options as follows:

```
proc nlin data=clarke1987 bias hougaard;
  parms theta1=-0.15 theta2=2.0 theta3=0.80;
  model y = theta3 + theta2*exp(theta1*x);
run;
```

Figure 4 displays the estimated bias and skewness in addition to the parameter estimates. The guideline, according to Ratkowsky (1983), is that percentage bias greater than 1% is considered to be significantly nonlinear. Similarly, a value of the standardized Hougaard skewness measure greater than 0.25 in absolute value indicates nonlinear behavior. Based on these guidelines, the estimates of both  $\theta_2$  and  $\theta_3$  are significantly biased and skewed, whereas the bias and skewness of estimate of  $\theta_1$  are negligible. This makes the application of standard inference to  $\theta_2$  and  $\theta_3$  unreliable, and it conflicts with the intuitive expectation that the estimates for  $\theta_1$  would show the strongest effects of nonlinearity.

**Figure 4** Box's Bias and Hougaard's Skewness Measures  
**The NLIN Procedure**

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits		Skewness	Bias	Percent Bias
theta1	-0.1031	0.0255	-0.1599	-0.0462	-0.0619	-0.00026	0.26
theta2	2.5190	0.2658	1.9268	3.1112	1.5428	0.0802	3.18
theta3	0.9631	0.3216	0.2466	1.6797	-1.3956	-0.0772	-8.02

### Impact of Nonlinearity on Parameter Uncertainty: The PROFILE Statement

Table 1 summarizes the options available in the PROFILE statement.<sup>2</sup>

**Table 1** Summary of Options in PROFILE Statement

Option	Description
ALL	Produces all profiling-related plots and statistics
CONFCURV	Produces confidence curves for the listed parameters
JACKKNIFE	Produces plots of the absolute relative percentage change by observation for the listed parameters
PROFDATA=	Specifies the profile output data set
RANGE=	Specifies a profiling range
TPLOT	Produces profile $t$ plots for the listed parameters

The impact of nonlinearity on the uncertainty of individual parameters can be efficiently summarized by parameter profiling (Bates and Watts 1988; Cook and Weisberg 1990). Parameter profiling consists of constraining a specified profile parameter,  $\theta$ , at a range of values around the original least squares estimate and performing least squares minimization on the rest of the parameters,  $\Theta$ . The results are collected to produce the likelihood ratio pivotal statistic,  $\tau(\theta)$ ,

$$\tau(\theta) = \text{sign}(\theta - \hat{\theta}) \left( \frac{\text{SSE}(\theta, \tilde{\Theta}) - \text{SSE}(\hat{\theta}, \hat{\Theta})}{\text{MSE}} \right)^{1/2}$$

<sup>2</sup>For information about these options, see the NLIN procedure documentation. This paper concentrates on diagnosing nonlinearity and reparameterization of nonlinear regression models. Only the options highly relevant to this purpose are discussed.

and the Wald pivotal statistic,  $\sigma(\theta)$ ,

$$\sigma(\theta) = \frac{\theta - \hat{\theta}}{\text{StdErr}_{\theta}}$$

where  $\theta$ ,  $\hat{\theta}$ , and  $\text{StdErr}_{\theta}$  are, respectively, the constrained value, the original least squares estimate, and the standard error for the profile parameter; MSE is the mean square error;  $\text{SSE}(\theta, \hat{\Theta})$  is the least squares sum of squares error with  $\theta$  constrained at a certain value; and  $\text{SSE}(\hat{\theta}, \hat{\Theta})$  is the original least squares sum of squares error. For linear models,  $\tau(\theta)$  matches  $\sigma(\theta)$ .

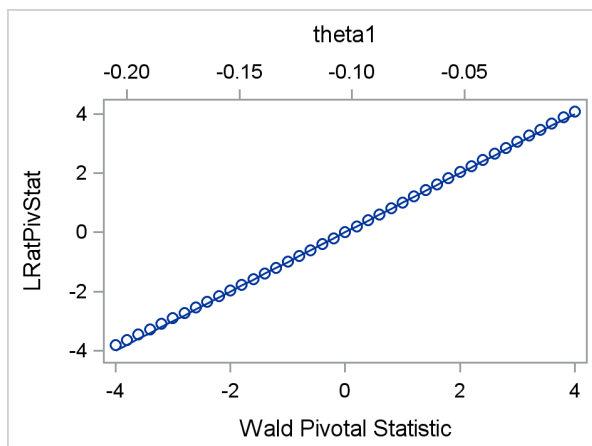
A profile  $t$  plot for parameter  $\theta$  is a plot of the likelihood ratio pivotal statistic,  $\tau(\theta)$ , versus the Wald pivotal statistic,  $\sigma(\theta)$ . In contrast, a confidence curve contains a scatter plot of the constrained parameter values versus  $\text{sign}(\theta - \hat{\theta}) \tau(\theta)$ . The standard confidence intervals are overlaid as two straight lines that pass through  $(0, \hat{\theta})$  with a slope of  $\pm \text{StdErr}_{\theta}$ . Hence, for different levels of significance, you can easily compare the standard confidence intervals to the corresponding confidence intervals that are based on the likelihood ratio (Cook and Weisberg 1990).

The following statements include the PROFILE statement to generate the profile  $t$  plots and confidence curves of the parameters:<sup>3</sup>

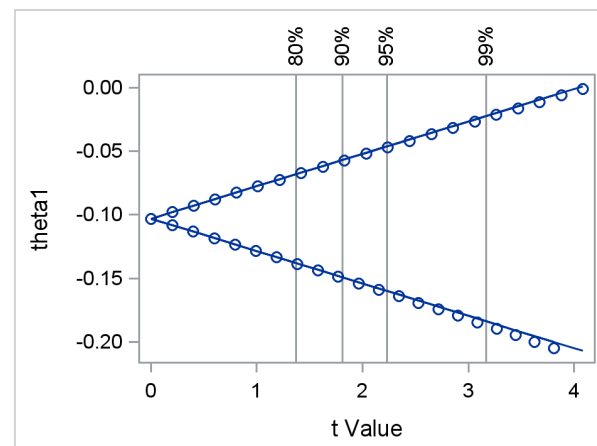
```
ods graphics on;
proc nlin data=clarke1987 bias hougaard;
  parms theta1=-0.15 theta2=2.0 theta3=0.80;
  model y = theta3 + theta2*exp(theta1*x);
  profile theta1 / range = -4 to 4 by 0.2 all;
  profile theta2 theta3 / range = -2 to 2 by 0.2 all;
run;
```

The profile  $t$  plots (Figure 5, Figure 7, and Figure 9) and confidence curves (Figure 6, Figure 8, and Figure 10) of the parameters corroborate the conclusions from Box's bias and Hougaard's skewness measures. Specifically,  $\theta_1$  behaves so linearly that the likelihood ratio statistic,  $\tau(\theta_1)$ , and the Wald pivotal statistic,  $\sigma(\theta_1)$ , are practically indistinguishable. As for  $\theta_2$  and  $\theta_3$ , there is a strong discrepancy between the two statistics that invalidates the standard inference. It is recommended that you avoid reporting a single confidence interval based on the standard inference unless there is good agreement between the standard interval and the likelihood intervals up to at least the 99% confidence level (Cook and Weisberg 1990).  $\theta_1$  satisfies this criterion, but  $\theta_2$  and  $\theta_3$  violate it strongly. This contradicts the intuitive expectation.

**Figure 5** Profile  $t$  Plot of  $\theta_1$

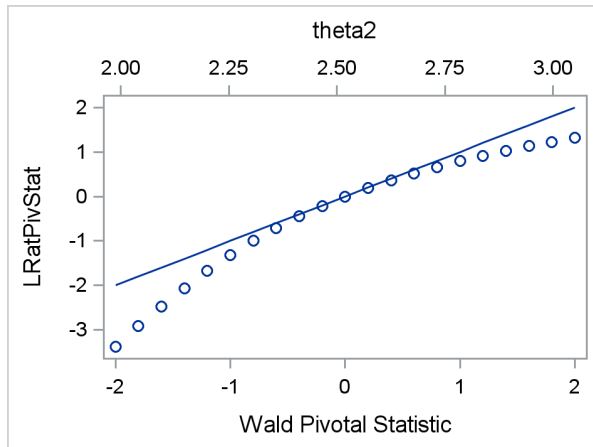


**Figure 6** Confidence Curve of  $\theta_1$

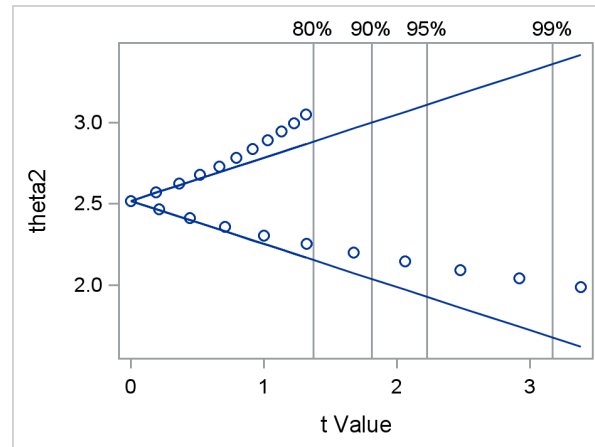


<sup>3</sup>The profile range of  $\theta_1$  is different from that of  $\theta_2$  and  $\theta_3$  so that the same range of significance levels in all three confidence curves is covered. Also, the titles from the individual profile  $t$  plot and confidence curves have been turned off to conserve space.

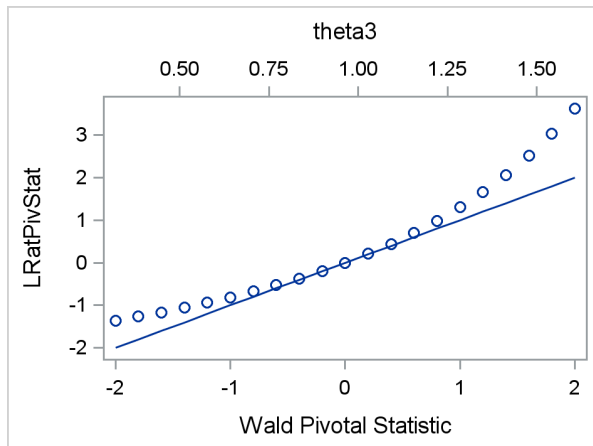
**Figure 7** Profile  $t$  Plot of  $\theta_2$



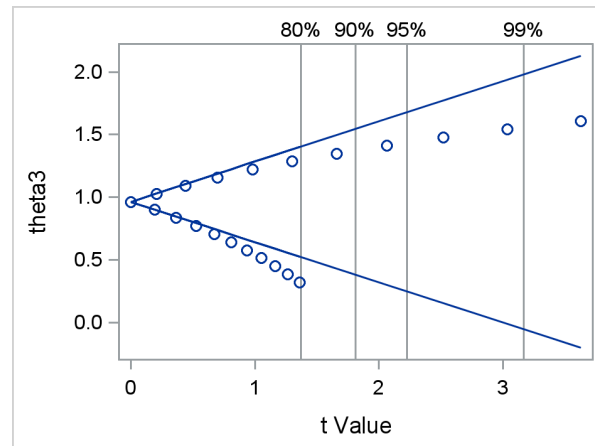
**Figure 8** Confidence Curve of  $\theta_2$



**Figure 9** Profile  $t$  Plot of  $\theta_3$



**Figure 10** Confidence Curve of  $\theta_3$



**Distribution of the Parameter Estimates: The BOOTSTRAP Statement**

Table 2 summarizes the options available in the BOOTSTRAP statement.<sup>4</sup>

**Table 2** Summary of Options in BOOTSTRAP Statement

Option	Description
BOOTCI	Produces bootstrap confidence intervals of the parameters
BOOTCORR	Produces a bootstrap correlation matrix estimate table
BOOTCOV	Produces a bootstrap covariance matrix estimate table
BOOTDATA=	Specifies the bootstrap output data set
BOOTPLOTS	Produces plots of the bootstrap parameter estimates
DGP=	Specifies the bootstrap data generating process (DGP)
NSAMPLES=	Specifies the number of bootstrap sample data sets (replicates)
SEED=	Provides the seed that initializes the random number stream

The data generating process (DGP) that the NLIN procedure uses during the bootstrap process can be represented by

$$\tilde{Y}_i = f(\hat{\theta}; \mathbf{z}'_i) + \tilde{\epsilon}_i$$

<sup>4</sup>For information about these options, see the NLIN procedure documentation. This paper concentrates on diagnosing nonlinearity and reparameterization of nonlinear regression models. Only the options highly relevant to this purpose are discussed.

where  $\tilde{Y}_i$  is the  $i$ th simulated response,  $\hat{\theta}$  is the original least squares parameter estimate,  $\mathbf{z}'_i$  is the  $i$ th regressor vector, and  $\tilde{\epsilon}_i$  is the  $i$ th simulated error.

PROC NLIN makes several bootstrap DGP types available in the DGP option in the BOOTSTRAP statement. The different bootstrap DGPs differ only in how they obtain the  $i$ th simulated error,  $\tilde{\epsilon}_i$ . Succinctly,  $\tilde{\epsilon}_i$  can be represented as

$$\tilde{\epsilon}_i = s_r \hat{\epsilon}_r$$

where  $r$  is a uniform random integer between 1 and the number of usable observations,  $s_r$  is a scale factor that depends on the chosen bootstrap DGP option, and  $\hat{\epsilon}_r$  is the  $r$ th residual obtained from the  $r$ th usable observation of the original least squares fit.

The following statements invoke the BOOTSTRAP statement to produce the bias-corrected bootstrap confidence intervals of the parameter estimates (DiCiccio and Efron 1996) and generate histograms and pairwise scatter plots of the parameter estimates:<sup>5</sup>

```
ods graphics on;
proc nlin data=clarke1987;
  parms theta1=-0.15 theta2=2.0 theta3=0.80;
  model y = theta3 + theta2*exp(theta1*x);
  bootstrap / seed=123 nsamples=5000 bootci bootplots(all);
run;
```

Figure 11 shows that the BOOTCI option of the BOOTSTRAP statement includes columns that contain the bias-corrected bootstrap confidence intervals in the parameter estimates table. The 95% standard confidence interval of  $\theta_1$ ,  $[-0.1599, -0.0462]$ , tightly covers the corresponding bias-corrected confidence interval,  $[-0.1508, -0.0535]$ . In contrast, the 95% standard confidence intervals of  $\theta_2$ ,  $[1.9268, 3.1112]$ , and  $\theta_3$ ,  $[0.2466, 1.6797]$ , have only loose and lower coverage of their respective bias-corrected confidence intervals,  $[2.1838, 3.5162]$  and  $[-0.2531, 1.3731]$ . Both in theory and in practice, the coverage of bootstrap-based confidence intervals is usually very close to the nominal value (DiCiccio and Efron 1996). Thus, taking the bias-corrected bootstrap confidence interval as a reference, you can conclude that the standard confidence intervals for  $\theta_2$  and  $\theta_3$  are unreliable.

**Figure 11** Parameter Estimates, Standard Confidence Intervals, and Bootstrap Confidence Intervals

### The NLIN Procedure

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits		Bootstrap Std Dev	Bootstrap Bias-Corrected 95% Confidence Limits	
theta1	-0.1031	0.0255	-0.1599	-0.0462	0.0250	-0.1508	-0.0535
theta2	2.5190	0.2658	1.9268	3.1112	0.4123	2.1838	3.5162
theta3	0.9631	0.3216	0.2466	1.6797	0.4608	-0.2531	1.3731

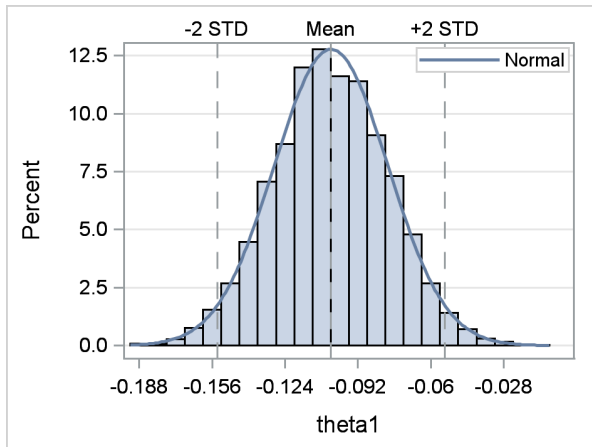
The histogram of estimates of  $\theta_1$  (Figure 12) closely follows the normal distribution. In contrast, the histogram of  $\theta_2$  (Figure 14) has a long right tail, and the histogram of  $\theta_3$  (Figure 16) has a long left tail. In addition, the pairwise scatter plots (Figure 13, Figure 15, and Figure 17) show strong deviations from the elliptical shape of pairwise scatter plots of normally distributed random variables.

These findings support the earlier assertion, based on Box's bias, Hougaard's skewness, and the parameter profiling, that  $\theta_2$  and  $\theta_3$  show significant nonlinear behavior but  $\theta_1$  does not. The section "Bootstrap-Based Reparameterization" on page 10 explains how you can use distribution of the parameter estimates to gain insight into a reparameterization that can reduce the parameter-effects curvature and bring the model closer to linearity.

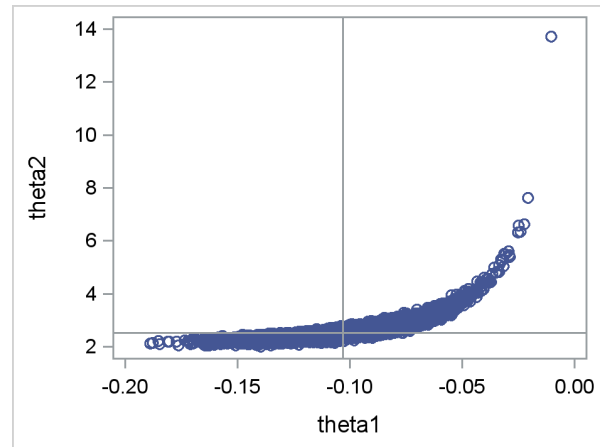
<sup>5</sup>The titles from the individual histogram and pairwise scatter plots have been turned off to conserve space.



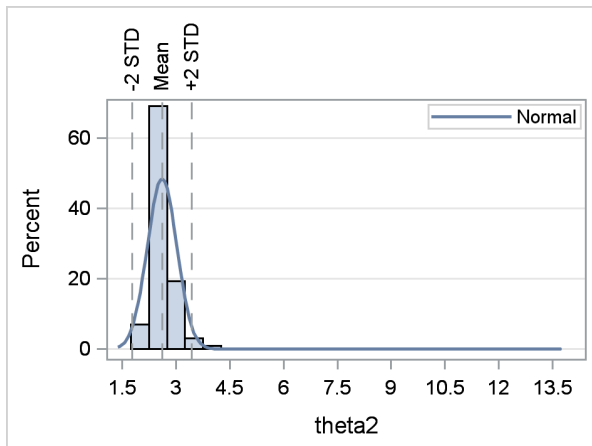
**Figure 12** Histogram of Estimates of  $\theta_1$



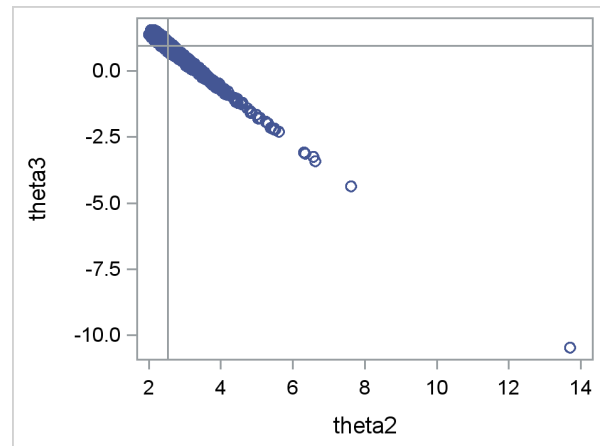
**Figure 13** Scatter Plot of Estimates of  $\theta_2$  by  $\theta_1$



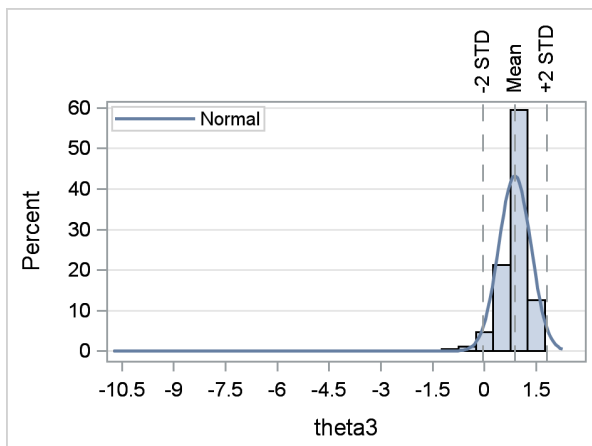
**Figure 14** Histogram of Estimates of  $\theta_2$



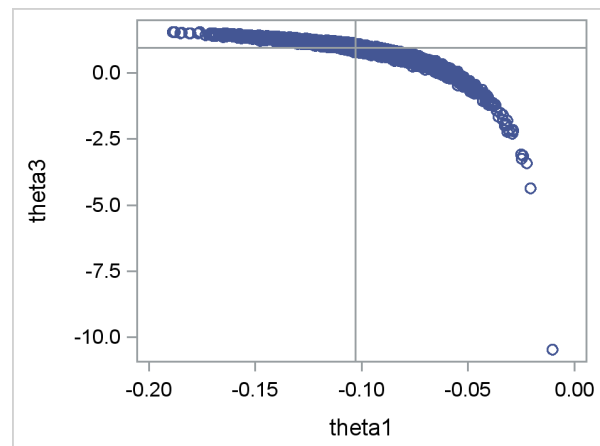
**Figure 15** Scatter Plot of Estimates of  $\theta_3$  by  $\theta_2$



**Figure 16** Histogram of Estimates of  $\theta_3$



**Figure 17** Scatter Plot of Estimates of  $\theta_3$  by  $\theta_1$



## DEMONSTRATION OF REPARAMETERIZATION TECHNIQUES

The ultimate objective of diagnosing a far-from-linear nonlinear model is to reparameterize the model so that it becomes a close-to-linear model. This is possible only for nonlinear regression models that have insignificant intrinsic curvature. Fortunately, most nonlinear regression models that are used in practice seem to have very small relative intrinsic curvatures (Bates and Watts 1980; Ratkowsky 1990).

The first step in reparameterization is to identify the reparameterization parameters of interest. For the nonlinear regression model at hand, the bias and skewness measures, the parameter profiling, and the bootstrap consistently identify  $\theta_2$  and  $\theta_3$  as parameters that show strong nonlinear behavior. Consequently, these two parameters are the reparameterization parameters of interest. After you identify the parameters of interest, you can use several different approaches to systematically reparameterize a nonlinear regression model (Hougaard 1982; Ratkowsky 1990). The following sections demonstrate two of these approaches: bootstrap-based reparameterization and expected-value reparameterization.

### Bootstrap-Based Reparameterization

Distribution of estimates of the parameters, produced by the bootstrap method and depicted as histograms, can be used to devise a reparameterization. This is especially true if the resulting distributions are long-tailed. If a parameter estimate's distribution has a long right tail, indicating that it is closer to a lognormal distribution than to a normal distribution, Ratkowsky (1990) recommends replacing the old parameter with its exponential or positive power. If the distribution has a long left tail, replacing the old parameter with its logarithm or fractional power is recommended.

The distribution of estimates of  $\theta_2$  (Figure 14) has a long right tail. Thus  $\theta_2$  is replaced by an exponential of a new parameter,  $\beta_2$ ,

$$\theta_2 = \exp(\beta_2)$$

In contrast, the distribution of estimates of  $\theta_3$  (Figure 16) has a long left tail. Hence, it is replaced by a logarithm of a new parameter,  $\beta_3$ ,

$$\theta_3 = \log(\beta_3)$$

The newly parameterized model expression is given by

$$y = \log(\beta_3) + \exp(\beta_2 + \theta_1 x)$$

The following statements fit this newly parameterized model, with the starting parameter values adjusted accordingly:

```
proc nlin data=clarke1987 nlinmeasures;
  parms theta1 = -0.15 beta2 = 0.69 beta3 = 2.23;
  theta2 = exp(beta2);
  theta3 = log(beta3);
  model y = theta3 + theta2*exp(theta1*x);
run;
```

Figure 18 shows the nonlinearity measures table for the new parameterization. Both the maximum and root mean square relative parameter-effects curvatures (2.6825 and 1.2308, respectively) are only about a third of their original values in Figure 3 (7.0183 and 3.1301, respectively). As expected, the relative intrinsic curvature measures do not change.<sup>6</sup> Even with these significant drops in their values, the new maximum and root mean square relative parameter-effects curvature are far above the critical curvature value, 0.5193. Although the new parameterization failed to make the model a close-to-linear model, it has significantly reduced the relative parameter-effects curvature. As a result, the newly parameterized model can be deemed to have better inferential and predictive properties.

---

<sup>6</sup>Of course, the residual variances do not change, either.

**Figure 18** Global Measures of Nonlinearity  
**The NLIN Procedure**

Global Nonlinearity Measures	
Max Intrinsic Curvature	0.0935
RMS Intrinsic Curvature	0.0418
Max Parameter-Effects Curvature	2.6825
RMS Parameter-Effects Curvature	1.2308
Curvature Critical Value	0.5193
Raw Residual Variance	0.0053
Projected Residual Variance	0.0059

### Expected-Value Reparameterization

Let  $m$  be the number of reparameterization parameters of interest. Expected-value reparameterization involves the following steps:

1. Choose  $m$  regressor vectors,  $\mathbf{z}_1, \dots, \mathbf{z}_m$ , from the observation.
2. Set the values of  $\mathbf{z}$  in the mean function,  $f(\boldsymbol{\theta}; \mathbf{z})$ , for each of the  $m$  regressor vectors and generate  $m$  equations of the form  $E[Y_i] \equiv y_i = f(\boldsymbol{\theta}; \mathbf{z}_i)$ . Do not replace  $\boldsymbol{\theta}$  and  $E[Y_i]$  with their numeric values.
3. Analytically solve the resulting simultaneous equations for the reparameterization parameters of interest in terms of  $y_i$  and the parameters that are not to be reparameterized.

Expected-value reparameterization is so called because the new parameters that result from the preceding steps are the expected values,  $y_i$ , for the chosen regressor values,  $\mathbf{z}_i$ . If you can complete these steps, in most cases this approach succeeds in producing a close-to-linear model (Ratkowsky 1990). However, lack of an analytical solution for the expected-value parameters, in the third step, can limit the universal applicability of expected-value reparameterization.

Because the reparameterization parameters of interest for the nonlinear regression at hand are  $\theta_2$  and  $\theta_3$ , the expected-value reparameterization can be applied by choosing two regressor values,  $\mathbf{z}_1 \equiv x_1$  and  $\mathbf{z}_2 \equiv x_2$ . The resulting simultaneous equation can be solved for  $\theta_2$  and  $\theta_3$  as

$$\theta_2 = \frac{\beta_3 - \beta_2}{\exp(\theta_1 x_2) - \exp(\theta_1 x_1)}$$

$$\theta_3 = \frac{\beta_2}{1 - \exp(\theta_1(x_1 - x_2))} - \frac{\beta_3}{\exp(\theta_1(x_2 - x_1)) - 1}$$

where, for convenience,  $\beta_2$  and  $\beta_3$  stand for the expected value parameters,  $y_1 \equiv E(Y|x_1)$  and  $y_2 \equiv E(Y|x_2)$ , respectively.<sup>7</sup>

Setting  $x_1$  and  $x_2$  to the minimum and maximum values of the regressor,  $x$ , the following statements fit this newly parameterized model with parameters  $\theta_1$ ,  $\beta_2$ , and  $\beta_3$ , and with starting values of the parameters adjusted accordingly:

```
proc nlin data=clarke1987 nlinmeasures bias hougaard;
  parms theta1 = -0.15 beta2 = 2.52 beta3 = 1.09;
  x1 = 1;
  x2 = 13;
  theta2 = (beta3 - beta2) / (exp(theta1*x2) - exp(theta1*x1));
```

<sup>7</sup>For the sake of brevity, the model expression, in terms of the new parameter set  $\theta_1$ ,  $\beta_2$ , and  $\beta_3$ , is omitted.

```

theta3 = beta2/(1 - exp(theta1*(x1 - x2))) -
        beta3/(exp(theta1*(x2 - x1)) - 1);
model y = theta3 + theta2*exp(theta1*x);
profile theta1 beta2 beta3 / range = -4 to 4 by 0.2 all;
bootstrap / seed=123 nsamples=5000 bootci bootplots(all);
run;

```

Figure 19 shows that the new parameterization succeeds in making the nonlinear regression model a close-to-linear model. This is supported by the relative intrinsic and parameter-effects curvature values (0.0935, 0.0418, 0.2080, 0.0949) being significantly less than the critical curvature value (0.5193). Further manifestations of the new close-to-linear model include the following:

- The percentage bias and skewness of the parameter estimates are significantly less than the bias and skewness threshold values, 1% and 0.25, that can be used to deem parameters biased and skewed.
- The 95% standard confidence intervals of all parameters (not only  $\theta_1$ ) show very tight coverage of the bias-corrected bootstrap confidence intervals.
- The profile  $t$  plots (Figure 21, Figure 23, and Figure 25) and confidence curves (Figure 22, Figure 24, and Figure 26) of the parameters show practically no difference between the likelihood ratio pivotal statistic,  $\tau(\theta)$ , and the Wald pivotal statistic,  $\sigma(\theta)$ , over the wide profiling range.
- The histograms of  $\beta_2$  (Figure 29) and  $\beta_3$  (Figure 31) are notably closer to the normal curve than are the long-tailed histograms of  $\theta_2$  (Figure 14) and  $\theta_3$  (Figure 16). This is in spite of the slight deviations of the histograms of  $\beta_2$  and  $\beta_3$  from the normal curve, compared to that of  $\theta_1$  (Figure 27).
- The pairwise scatter plots of the parameters (Figure 28, Figure 30, and Figure 32) are elliptical in shape, matching the expectation from the pairwise scatter plot of two normally distributed random variables. These are in clear contrast with the corresponding plots (Figure 13, Figure 15, and Figure 17) of the original far-from-linear nonlinear model, which have curled and elongated shapes.

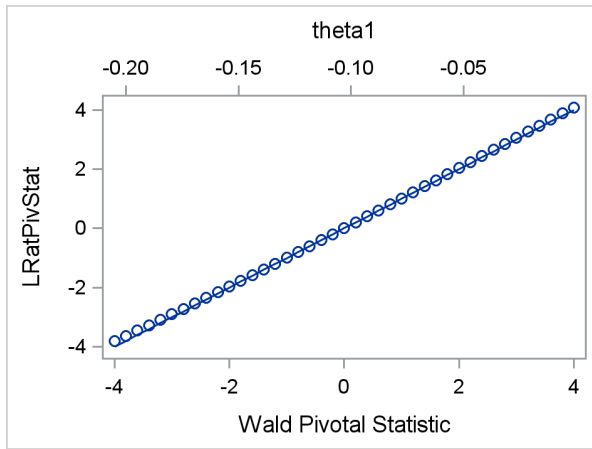
**Figure 19** Global Measures of Nonlinearity  
**The NLIN Procedure**

Global Nonlinearity Measures	
Max Intrinsic Curvature	0.0935
RMS Intrinsic Curvature	0.0418
Max Parameter-Effects Curvature	0.2080
RMS Parameter-Effects Curvature	0.0949
Curvature Critical Value	0.5193
Raw Residual Variance	0.0053
Projected Residual Variance	0.0059

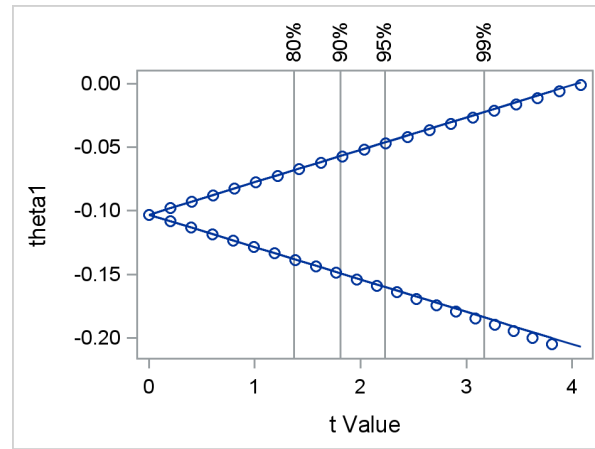
**Figure 20** Parameter Estimates, Standard Confidence Intervals, and Bootstrap Confidence Intervals

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits		Bootstrap Std Dev	Bootstrap Bias-Corrected 95% Confidence Limits		Skewness	Bias	Percent Bias
theta1	-0.1031	0.0255	-0.1599	-0.0462	0.0250	-0.1508	-0.0535	-0.0619	-0.00026	0.26
beta2	3.2355	0.0574	3.1077	3.3632	0.0564	3.1081	3.3242	0.0378	0.000860	0.027
beta3	1.6229	0.0473	1.5174	1.7284	0.0469	1.5154	1.7006	-0.0585	-0.00092	-0.06

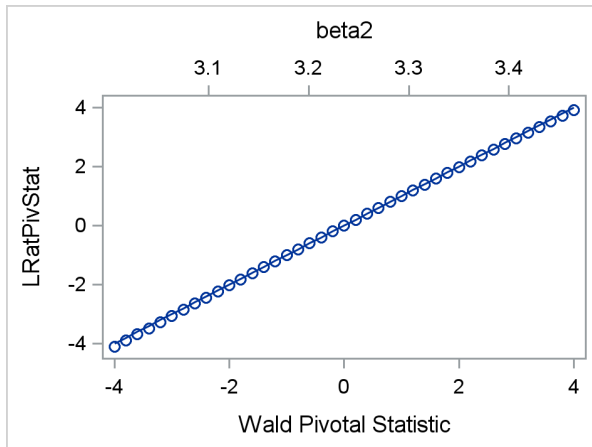
**Figure 21** Profile  $t$  Plot of  $\theta_1$



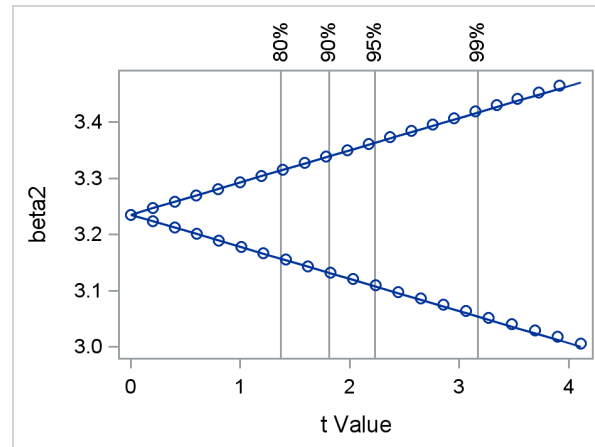
**Figure 22** Confidence Curve of  $\theta_1$



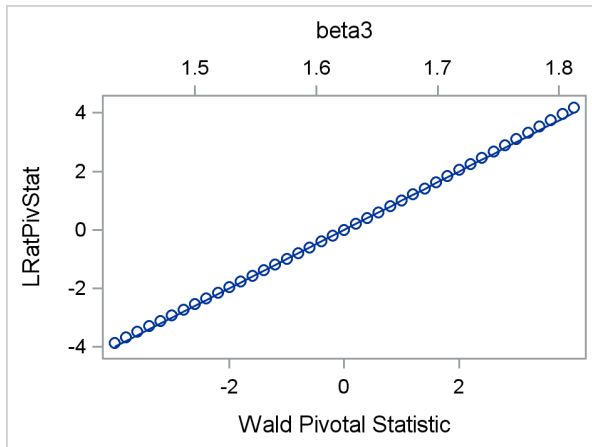
**Figure 23** Profile  $t$  Plot of  $\beta_2$



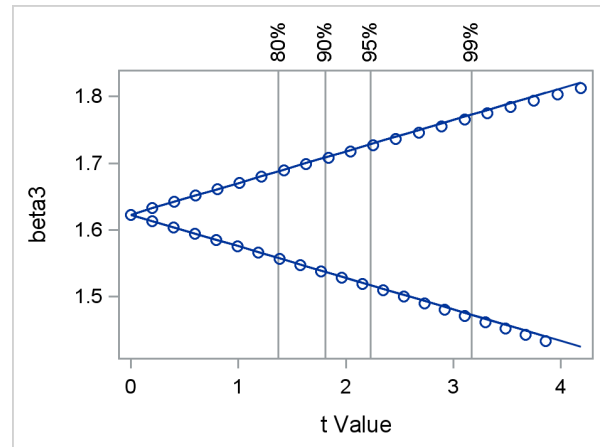
**Figure 24** Confidence Curve of  $\beta_2$



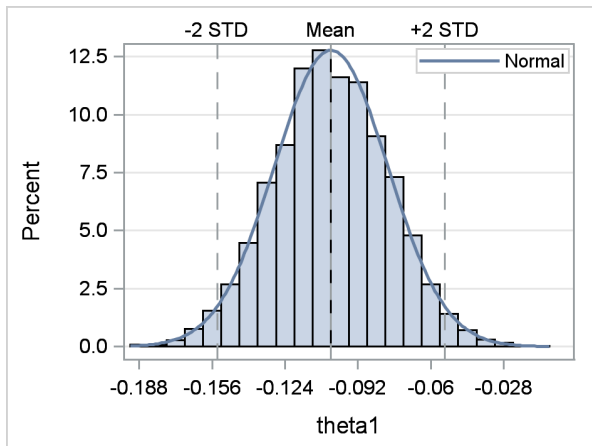
**Figure 25** Profile  $t$  Plot of  $\beta_3$



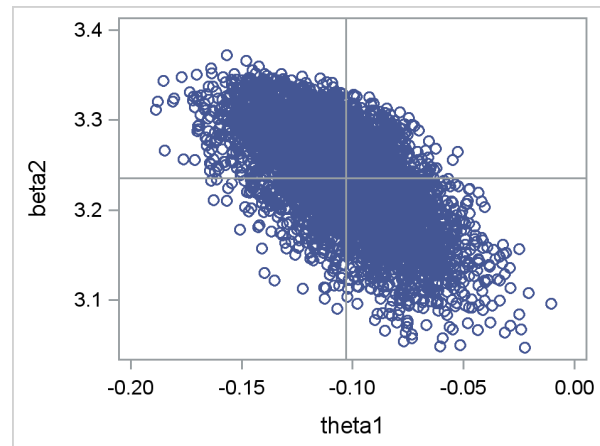
**Figure 26** Confidence Curve of  $\beta_3$



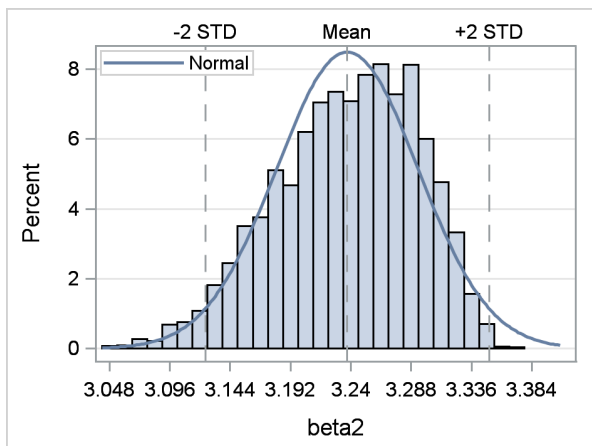
**Figure 27** Histogram of Estimates of  $\theta_1$



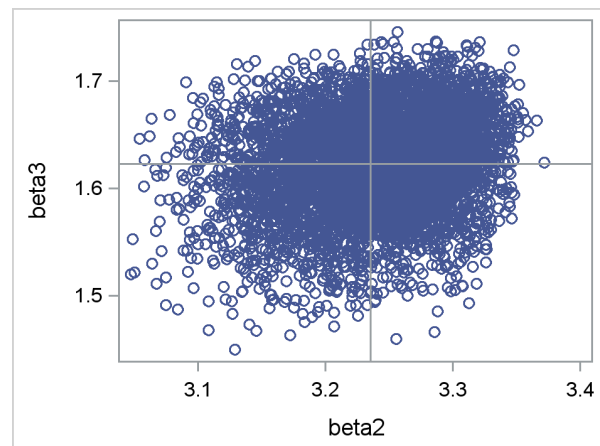
**Figure 28** Scatter Plot of Estimates of  $\beta_2$  by  $\theta_1$



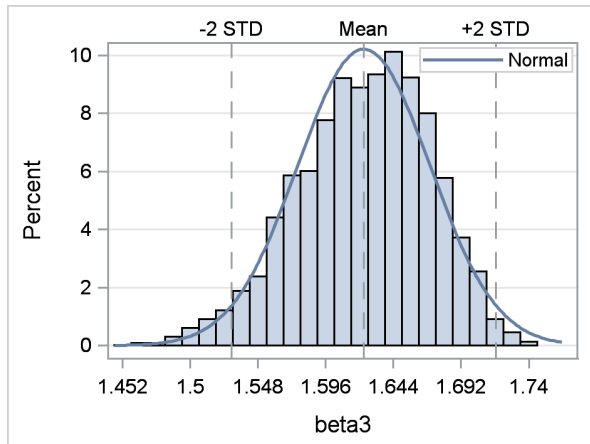
**Figure 29** Histogram of Estimates of  $\beta_2$



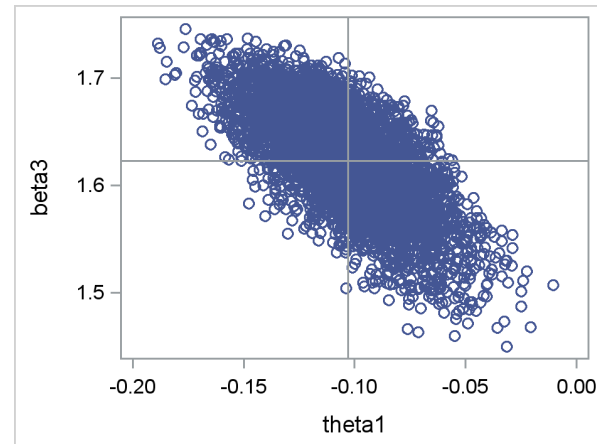
**Figure 30** Scatter Plot of Estimates of  $\beta_3$  by  $\beta_2$



**Figure 31** Histogram of Estimates of  $\beta_3$



**Figure 32** Scatter Plot of Estimates of  $\beta_3$  by  $\theta_1$



## MULTITHREADING IN THE NLIN PROCEDURE

The BOOTSTRAP and PROFILE statements in PROC NLIN result in a large number of independent nonlinear least squares minimizations. For the BOOTSTRAP statement, the number of minimizations that must be carried out is equal to the value of the NSAMPLES option in the BOOTSTRAP statement. The number of independent minimizations that the PROFILE statement generates depends on the number of parameters to be profiled and the number of constrained values in the specified profiling range (of each profile parameter).

The independence of these minimization tasks is very favorable for parallelization on modern multicore machines. Even though, in theory, both data-based parallelization and task-based parallelization are possible, the NLIN procedure uses task-based parallelization, simultaneously working on as many minimization problems as there are CPU cores. This approach helps the procedure efficiently use all available CPU cores and eliminates most of the reproducibility issues that can arise in data-based parallelization.

## SUMMARY AND CONCLUSION

The NLIN procedure added several diagnostic and inferential features, such as the relative measures of curvature, Box's bias, the bootstrap, and parameter profiling, in SAS/STAT 9.3, 12.1, and 13.1. Focusing on nonlinearity and its effects on statistical inference and prediction, this paper demonstrates how you can use these features to progressively diagnose and reparameterize a far-from-linear nonlinear regression model.

The NLINMEASURES option in the PROC NLIN statement produces global measures of nonlinearity that can diagnose a far-from-linear nonlinear regression model. Box's bias and Hougaard's skewness measures, produced by the BIAS and HOUGAARD options in the PROC NLIN statement, estimate the bias and skewness of parameter estimates. The bias and skewness estimates can identify the parameters that show strong nonlinear behavior. The PROFILE and BOOTSTRAP statements, which were added in SAS/STAT 12.1 and 13.1, respectively, can be used to provide a more extensive understanding of the inferential and distributional properties of the parameter estimates in a nonlinear regression model. The paper also demonstrates approaches to reparameterization that are based on the distribution of the bootstrap parameter estimates and the technique of expected-value parameterization.

In conclusion, the NLIN procedure, with a host of added features, enables you to make better statistical inference. In situations where a nonlinear regression model needs reparameterization, you can also use these features to iteratively perform and gauge reparameterizations, progressively improving the inferential and predictive properties of the model. Furthermore, by using task-based parallelization, the NLIN procedure tames the plethora of independent minimizations that have to be performed during parameter profiling and the bootstrap. Hence, the oft-stated reason to shun these features—computational cost—should be less of a concern.

## REFERENCES

- Bates, D. M. and Watts, D. G. (1980), "Relative Curvature Measures of Nonlinearity (with Discussion)," *Journal of the Royal Statistical Society, Series B*, 42, 1–25.
- Bates, D. M. and Watts, D. G. (1988), *Nonlinear Regression Analysis and Its Applications*, New York: John Wiley & Sons.
- Box, M. J. (1971), "Bias in Nonlinear Estimation (with Discussion)," *Journal of the Royal Statistical Society, Series B*, 33, 171–201.
- Clarke, G. P. Y. (1987), "Approximate Confidence Limits for a Parameter Function in Nonlinear Regression," *Journal of the American Statistical Association*, 82, 221–230.
- Cook, R. D. and Weisberg, S. (1990), "Confidence Curves in Nonlinear Regression," *Journal of the American Statistical Association*, 85, 544–551.
- DiCiccio, T. J. and Efron, B. (1996), "Bootstrap Confidence Intervals," *Statistical Science*, 11, 189–212.
- Hougaard, P. (1982), "Parameterizations of Nonlinear Models," *Journal of the Royal Statistical Society, Series B*, 44, 244–252.
- Ratkowsky, D. (1983), *Nonlinear Regression Modeling*, New York: Marcel Dekker.
- Ratkowsky, D. (1990), *Handbook of Nonlinear Regression Models*, New York: Marcel Dekker.
- Seber, G. A. F. and Wild, C. J. (1989), *Nonlinear Regression*, New York: John Wiley & Sons.

## ACKNOWLEDGMENTS

The author is grateful to Yiu-Fai Yung and Ed Huddleston of the Advanced Analytics Division at SAS Institute Inc. for their valuable assistance in the preparation of this manuscript.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author:

Biruk Gebremariam  
SAS Institute Inc.  
SAS Campus Drive  
Cary, NC 27513  
919-531-6271 (phone)  
919-677-4444 (fax)  
biruk.gebremariam@sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.