THE INSTITUTE OF STATISTICS

UNIVERSITY OF NORTH CAROLINA SYSTEM

MODEL SELECTION AND PARAMETER ESTIMATION FOR CENSORED DATA
FROM HIGHLY FRACTIONATED EXPERIMENTS

by

J. C. Lu and C. Unal

Institution of Statistics Mimeograph Series No. 2211

January, 1992

NORTH CAROLINA STATE UNIVERSITY
Raleigh, North Carolina

Model Selection and Parameter Estimation for Censored Data From Highly Fractionated Experiments

Jye-Chyi Lu and Cemal Unal
Department of Statistics
North Carolina State University
Raleigh, NC 27695-8203

Key Words – Reliability Improvement, Censored Data, Model Selection, Screening Experiments.

Reader Aids –

Purpose: Advance state of the art

Special math needed: Maximum likelihood estimation

Special math needed for results: None

Results useful to: Reliability analyses and statisticians

Abstract – Complicated structure of screening experiments coupled with censored data resulting from life-testing of durable products makes the characterization and optimization of process variables very difficult. Existing methods are inadequate for analyzing such data from highly fractionated experiments because important effects are identified wrongly. By studying contrasts from censored data, we can identify an appropriate initial model for the "imputation-modeling-maximization" algorithm to compare many models simultaneously and select the best model for process optimization. We demonstrate the procedure by reanalyzing data sets from Specht's heat exchanger experiments and show its superiority over some existing methods by comparing their predicted lifetimes.

1. INTRODUCTION

Conducting experimentations to improve quality/reliability of products becomes a routine practice in industries nowadays. However, due to the increasing durability of products, the data collected from life-testing experiments are usually censored or grouped. Incomplete information together with complexity of the design of experiment causes difficulties in s-modeling reliability measurements for improving the process control. In many models discussed in Section 3.2, the ML estimates do not exist are difficult to compute. Existing methods such as Taguchi's [14] minute accumulating analysis (MMA), the quick and dirty method mentioned by Hamada and Wu [7, pp. 26], and those proposed by Hahn-Morgan-Schmee [3] are inadequate for analyzing such censored data from highly fractionated experiments, because the best combination of levels of the process variables is suggested wrongly. Hamada and Wu [7] propose an iterative "expectation-modeling-maximization" (EMM) procedure to compare many models simultaneously. Their method takes the advantage of using existing software and promoting experimenter involvement. For brevity in future references, we refer to their paper as HW1.

Hamada and Wu's [7] EMM procedure considers the main-effect model as the initial model and obtains its ML estimates for imputing the censored data as pseudo-complete sample. Then, they select models based on the imputed data. It is very possible that the MLE in the main effect model does not exist and their procedure cannot be applied (c.f. HW1, pp 37). We propose to study all contrasts generated from main effects and interactions suggested from the structure of experimental design. By fitting the censored data one factor at a time, LIFEREG in the SAS [11] software can compute all these contrasts. We apply our procedure to select a model, for which the MLE exists. The "imputation-modeling-maximization" algorithm can then be used successively to select better models.

In analyzing complete samples resulting from screening experiments, Hamada and Wu [8] (denoted as HW2) propose a "significant-main-effect" (SME) model-selection procedure to identify important factors. They first identify s-significant main effect(s) and then screen out important interactions between the identified main effect(s) and the other factors. In our "mixed" model-selection method, we utilize the stepwise regression to compare all main effects and interactions. We do not limit our choice of model terms within the s-significant main effect(s) and its interactions with other main effects only. This method is demonstrated better than SME with several examples in Section 4.3.2.

Section 2 reviews the background of the problem and gives details about Specht's [13] heat exchanger life data. Statistical model of noises, transformation parameters and difficulties in the ML estimation are addressed. Hamada and Wu's [7] EMM procedure is reviewed in the end. In Section 3, we propose a model selection procedure and compare our method to Hamada and Wu's [8] EMM and SME procedures. Several examples are provided to show the improvement of our method. We also conduct a sensitivity study of the model selections under different quality of imputed data and possibly different types of parameter estimates. Section 4 and 5 give details of our selection of the final models and demonstrate their superiority over some existing analyses by comparing the predicted lifetimes. Section 6 give a brief conclusion of our study.

2. NOTATION

$(a,\ b)$	interval censored data
$(a,\ \infty)$	right-censored data
$\underset{\sim}{\beta}$	a vector of parameters for important effects
$\underset{\sim}{X}$	design matrix of process variables
$\phi(z),\ \Phi(z)$	the standard s-normal pdf and cdf, respectively

$\mu_i = \underset{\sim}{x}_i' \underset{\sim}{\beta}_i$ mean response at the ith run

$\hat{\underset{\sim}{\beta}}$, $\hat{\sigma}^2$ maximum likelihood estimates of $\underset{\sim}{\beta}$, σ^2 respectively

MLE maximum likelihood estimation

Other, standard notation is given in "Information for Readers & Authors" at the rear of each issue.

3. BACKGROUND OF THE PROBLEM AND LITERATURE REVIEW

3.1 The Heat Exchanger Life Data, Statistical Model and Box-Cox Transformation

To illustrate our methodology, we study the "wall" and "corner" data taken from Specht's [13] study of improving the reliability of a certain type of heat exchanger. In this article, we limit our discussion on the 12-run two-level Plackett-Burman design. The last second column of Table 1 lists the interval lifetimes (in 100 cycles) until the development of tube wall cracks, which is the data studied in HW1. The last column gives lifetimes till the development of tube corner cracks, where 5 out 12 data points are censored at 164. Using the notation in HW1, we assume that the random sample y after some transformation $h(y)$ follows a linear model with s-normal error:

$$h(y) = \underset{\sim}{X} \underset{\sim}{\beta} + \sigma \epsilon \quad \text{with } \epsilon \sim N(0, 1), \tag{3.1}$$

where $\underset{\sim}{\beta}$ is the vector of *important* effects and $\underset{\sim}{X}$ is the corresponding matrix of explanatory variables, including the main effects and interactions.

[Please place Table 1 here]

Considering the "wall" data (without multiplying the 100 cycles), Hamada and Wu [7] selected the final model (Intercept, E, EG, EH) for the data transformed (c.f. Box and Cox [1]) with $\lambda = -1.0$. When we search the best Box-Cox transformation parameter λ in fitting the s-normal distribution, for λ between -1.1 to -0.3, LIFEREG does not give converged results but the Powell's [10] program estimates the log-likelihood all at -2.772589. However, when λ changes from -0.2, -0.1, 0.0 ($\log_e$-

transformation), 0.1, 0.2, 0.5, to 1.0, the parameter estimates provided from LIFEREG are all converged and log-likelihoods are ranged from -5.858270 to -12.562295. For the "corner" data, since the main effects A and K are both s-significant at $\alpha = 0.05$ level, we entertain the model (Intercept, A, K) and plot (see Figure 1) log-likelihoods against transformation parameters λ from -1.0 to 1.0. Although the log-likelihood is minimized at $\lambda = -0.4$, we take $\lambda = -1.0$ for the future studies of both the "wall" and "corner" data for comparison purpose.

[Please place Figure 1 here]

3.2 Difficulties in the Maximum Likelihood Estimation

In analyzing censored data from highly fractionated experiments, there are estimability problems in the classical ML approach (see Silvapulle and Burridge [12]; Hamada and Tse [4]) because the paucity of data. Hamada and Tse [5] pointed out that certain design configuration coupled with a special censoring scheme will cause non-existence of the MLE. For example, in the 2^2 experimental plan, corresponding to run #1, #2, #3 and #4, the factors A and B are assigned as (+1, +1, -1, -1) and (+1, -1, +1, -1), respectively. They stated that if left- or right- censored data are contained in the first and second runs, denoted as 12, then the MLE's of the regression parameters and σ do not exist. In fact, if there are two censored data points, the MLE will exist only in cases that censoring occurs at 14 or 23 observations. Hamada and Tse [6] thus develop a MLECHK program based on a linear programming algorithm to check whether the MLE exists from censored data in a design of experiment. Besides using the MLECHK program, one can detect non-existence of the MLE problem from running LIFEREG in the SAS software, when the warning message of non-positive Hessian matrix is provided (c.f. Hamada and Tse [5], pp. 5). Our experience (c.f. HW1, pp. 37) indicates that in the case where the MLE does not exist, the likelihood surface is flat for a range of unbounded parameter values. Even when the MLE exists, in some cases, the likelihood can be almost the same for certain combinations of parameter values (see Example 1), which makes the search for the MLE difficult, and thus LIFEREG shows the nonconvergence message. In Lu and Unal [9], we propose a Bayesian model-selection and parameter-estimation procedure to overcome the drawback of the MLE. In next two examples, we examine the existence of the MLE, explore the likelihood around the ML estimate, and give a reliable procedure of getting the estimate.

Example 1. In fitting a regression model to the censored "wall" data, Hamada and Wu [7] gave the MLE $\hat{\beta}_0 = 0.984675$, $\hat{\beta}_E = -0.4252\text{e-}02$, $\hat{\beta}_{EG} = 0.2305\text{e-}02$, $\hat{\beta}_{EH} = -0.1927\text{e-}02$, $\hat{\sigma} = 1.02\text{e-}04$ and $\log_e$-likelihood = −2.772607. Although the use of Newton-Raphson procedure in LIFEREG does not give a converged result, the log-likelihood is reported as −8.56336 and the parameter estimate of $(\beta_0, \beta_E, \beta_{EG}, \beta_{EH}, \sigma)$ is given as (0.98347, −0.48027e-02, 0.20605e-02, −0.17918e-02, 1.36784e-04). To check whether the MLE in HW1's model exists, we first utilize Powell's [10] conjugate directions algorithm to evaluate numerous likelihoods across different combinations of parameter values. Treating HW1's MLE as the initial parameter value, we obtain the converged estimates as (0.984675, −0.425218E-02, 0.230502e-02, −0.192739e-02, 1.37747e-04) and the log-likelihood as −2.772589 after several iterations of the Powell's program itself. The estimates reported here are all s-significant to the digits changed in running FORTRAN programs. With the initial value given from LIFEREG, we obtain the estimates (0.984657, −0.426636e-02, 0.229084e-02, −0.190939e-02, 1.36459e-04) with the log-likelihood = −2.772589. Since these two estimates are slightly different, we further explore the likelihood surface around the estimates. We evaluate log-likelihoods of 3200 combinations of parameters with values β_0 = 0.984647, 0.984657, ..., 0.984687, β_E = −0.427636e-02, −0.426636e-02, ..., −0.424636e-02, β_{EG} = 0.228084e-02, 0.229084e-02, ..., 0.231184e-02, β_{EH} = −0.189939e-02, −0.190939e-02, ..., −0.193939e-02, and σ = 1.32459e-04, 1.33459e-04, ..., 1.39459e-04. We find that 70 combinations of parameters all give the highest log-likelihood at −2.772589 and their combinations are summarized in Table 2. From Table 2, we learn that these 70 cases are not randomly scattered but rather concentrated to a few combinations lined up together. Ninety combinations of parameters give log-likelihood -2.772590 and other cases studied all give log-likelihood under -2.779036. In conclusion, the flatness of the log-likelihood in certain region causes the trouble of using LIFEREG to find the MLE and the existence of the MLE indicated by the MLECHK program is correct. From this example, we learn that the unconverged estimate given from LIFEREG serve as a good starting point in using Powell's program for searching the ML estimate. □

[Please place Table 2 here]

Example 2. Fitting ten main effects to the "wall" data, we obtain a converged ML estimate with log-likelihood -12.540285 from LIFEREG. A study of the likelihood shows that it is not as flat as the case in Example 1. Fitting the same data set with (intercept, E, EG, EH, D, DH, DJ, BD) suggested in Example 5, we obtain the converged parameter estimates and its log-likelihood = -1.136512e-09. However, the MLE does not exist in that model with additional terms CD and DG. For the model (Intercept, E, D, CD, J, CJ, DH) selected from the contrasts (see Example 3), the MLE's exist up to the term J only. Considering the data set given from the "corner" measurements, we find that the MLE does not exist for the main-effect and some other models with two factor interactions (see Example 4). □

3.3 Hamada and Wu's [7] Model-Selection Method

Existing methods such as Taguchi's [14] minute accumulating analysis and Hahn-Morgan-Schmee's [3] procedure are inadequate for analyzing censored data from screening experiments, because the best combination of levels of process variables is suggested wrongly. Hamada and Wu [7] proposed the following "expectation-modeling-maximization" (EMM) algorithm to compare many models simultaneously. The key idea in EMM procedure is to solve the incomplete data problem by imputed pseudo-complete sample. This "imputation-maximization" technique is commonly used in reliability (c.f. Wei and Tanner [15]) and biostatistics (c.f. Dempster, Laird and Rubin, [2]). However, in papers other than HW1, their concentration is parameter estimation instead of selecting the best model and factor levels.

The EMM algorithm consists of the following steps: (A) Model selection phase: (A1) Initial model specification, (A2) Model fitting, (A3) Imputation, (A4) Model selection; repeat steps (A2) to (A4) until model selection termination; (B) Model assessment phase; repeat (A) and (B) until adequate model(s) are found; (C) Factor-level recommendation. In this article, we will present studies of the phase (A) steps only. Details of the procedure in part A are listed in the following paragraph and the marked (with $*$) steps will be modified in Section 4.

(A1) Initial Model Specification.* The experimenter decides the potentially important main effects, interactions and choose $\underset{\sim}{\mu}_0 = \underset{\sim}{X}_0 \underset{\sim}{\beta}_0$ (model 0). If no information is provided from the experimenter, they start with the model containing main effects only.

(A2) Model Fitting. Fit the current model $\underset{\sim}{\mu} = \underset{\sim}{X}_i \underset{\sim}{\beta}_i$, using the ML criterion. The likelihood of a complete sample is $\phi(z_y)\,|\partial h(y)/\partial y|$ and the likelihood of a censored data is $\Phi(z_b) - \Phi(z_a)$, where $z_w = [h(w) - \underset{\sim}{\mu}]/\sigma$ and h is the transformation in (3.1).

(A3) Imputation.* Impute the censored data by their conditional expectation,

$$\mathrm{E}[h(y) \mid y \in (a, b)] = \underset{\sim}{x}_i \underset{\sim}{\beta}_i + \sigma[\phi(z_a) - \phi(z_b)]/[\Phi(z_b) - \Phi(z_a)]. \quad (3.2)$$

(A4) Model Selection.* Informally apply a standard technique, such as stepwise regression, to identify s-significant main effect(s) and then screen out important interactions between the identified main effect(s) and the other factors. They stop the phase (A) when the current model selected is the same as the previous model.

Remark: In the initial model specification, the use of the main-effect model might not be feasible when the MLE does not exist. In next section, we propose a procedure to answer the question of how to start with the EMM algorithm, when the MLE's in selected initial models do not exist. Moreover, Hamada and Wu's SME model selection procedure is simple but can be deficient since it only considers the interactions identified by the s-significant main effects only. There might be many s-significant interactions that their main effects are not the most s-significant ones, see examples in next section. □

4. THE PROPOSED MODEL SELECTION PROCEDURE

4.1 Initial Model Selection

Instead of fitting the initial main-effect model, in our procedure, we start with computing contrasts of all factors one at a time by grouping the censored data in different levels according to the experimental layout (c.f. Box and Meyer [1]). For example, considering the factor E in the "wall" heat exchanger data, we computed the ML estimates of the mean and variance for observations #1, #3, #5, #9-11 as 0.988959 and 0.002605, respectively, and 0.980102, 0.003336 for the mean and variance of the other observations. The contrast of factor E is then estimated as -0.008858. Compared to the contrast of factor F estimated as 0.001001, the contrast of E is 8.85 times the one of F. Going through all these estimated contrasts, one can identify important main effects and interactions. For the convenience of screening all the contrasts, one can use LIFEREG to estimate regression coefficients with a single factor F, E, ..., AE, ..., etc. in the model and rank these factors according to their p-values. The use of this procedure assumes the equality of the variances of the observations in different levels, which is the basic assumption in the current study.

Example 3 (The initial model for the "wall" data). The order of the contrasts (or their p-values in LIFEREG) of 10 main effects and 45 interactions are given as follows:

$$\begin{aligned} & \mathrm{E} > \mathrm{CD} = \mathrm{CJ} = \mathrm{DH} > \mathrm{AJ} > \mathrm{GJ} > \mathrm{BK} > \mathrm{FD} > \mathrm{FB} = \mathrm{AK} \\ & \qquad > \mathrm{BG} = \mathrm{CK} = \mathrm{EG} = \mathrm{HK} > \cdots > \mathrm{EH} > \cdots, \end{aligned} \tag{4.1}$$

where E, EG and EH are the factors suggested by HW1. The p-values of the estimates of E, CD, BK, FD, FB, BG and EH are computed as 0.0001, 0.0251, 0.0498, 0.0564, 0.1128, 0.1261 and 0.1774, respectively. □

Example 4 (The initial model for the "corner" data). Ordering the contrasts computed from the censored "corner" data, we obtain

$$\begin{aligned} & \mathrm{A} = \mathrm{BG} > \mathrm{FC} > \mathrm{EJ} > \mathrm{K} = \mathrm{BD} > \mathrm{FD} = \mathrm{CG} > \mathrm{FJ} > \mathrm{AH} > \mathrm{EH} \\ & \quad > \mathrm{CJ} > \mathrm{D} = \mathrm{BK} = \mathrm{AE} = \mathrm{CH} = \mathrm{DJ} > \cdots. \end{aligned} \tag{4.2}$$

The p-values of the estimates of A, FC, K and FD are calculated as 0.0100, 0.0174, 0.0403, 0.1145, respectively. The effects A and K are both s-significant at the 0.05 level□

In Section 4.3, we present a procedure to select a model based on these orders of contrasts. Example 3 and 4 will be continued later.

4.2 Imputation

Instead of imputing the data at the conditional mean, one can randomly sample an observation from the truncated distribution $\phi(t)/[\Phi(z_b) - \Phi(z_a)]$, where $z_w = (t_w - \underset{\sim}{x}_i'\hat{\underset{\sim}{\beta}}_i)/\hat{\sigma}$, t_w is the transformed lower/upper interval point, $z_b = \infty$ for right-censored data, and $\underset{\sim}{x}_i'\hat{\underset{\sim}{\beta}}_i$ is the predicted mean at the ith run. This method works better than the conditional mean (3.2) for the "corner" data in Section 6.

4.3 Model Selection Procedure

Instead of using Hamada and Wu's [8] s-significant-main-effect method to select the model terms, we utilize stepwise regression to screen all main effects and interactions suggested by the design structure. It is possible that in stepwise regression some of two-factor interactions might be selected first without the presence of the main effects. To satisfy the heredity criterion (c.f. HW2), one needs to include the "induced" main effects first before entering the identified interaction. Hence, a "rank table" (*e.g.* Table 3) is constructed to organize the ranks of the effects and fill in the induced main effects with the frequencies of their usages in the interactions. Then, we can select a model with the least number of induced effects to increase the R-Square of the fitting. See next few examples for more details of rank tables and model-selections.

4.3.1 Model Selection Based on a Rank of Contrasts Calculated from Censored Data

Example 3 (Continued). Under the requirement of heredity, based on the ranks given in (4.1), if we select the interaction CD, we need to keep either the main effect C or D, where C has a larger contrast than D's. The factors C and D are called as the "induced" main effects. From the "rank table" (see Table 3), we select a model in the sequence of the factors entering into the model as follows:

$$\text{Intercept, E, D, CD, J, CJ, DH, AJ, GJ, B, BK,} \tag{4.3}$$

where we can keep 10 factors additional to the intercept and variance for these 12 data points. Because D is induced from CD and DH and J is induced from CJ, AJ and GJ, we keep D and J instead of keeping C, H and J to reduce the number of induced factors in the model. □

[Please place Table 3 here]

Example 4 (Continued). From the rank of contrasts (4.2), we select the following model as our initial model for the "corner" data:

$$\text{Intercept, A, K, B, BG, F, FC, E, EJ, BD, FD,} \tag{4.4}$$

We put K in the front due to its small p-value and the increment of log-likelihood. □

In next few examples, with some complete data, we compare our "mixed" model-selection procedure to Hamada and Wu's [8] "significant-main-effect" method.

4.3.2 *Comparison of Hamada and Wu's and Our Model Selection Procedures*

Example 5. In using the quick and dirty (QD) method (c.f. HW1) to analyze censored data, one imputes the interval data at their midpoints and treats the right-censored observation as the failure time, and applies regression methods to fit these pseudo-complete sample points. Imputing the "wall" data as the complete sample according to the QD method, in Table 4, we compare SME and our model selection procedures. From the report of R-Squares we conclude that our procedure works much better than HW2's. In our model,

$$\text{Intercept, E, B, BA, F, FA, BJ, FG, BC, J,} \tag{4.5}$$

the R-Squares jump from 0.7834 to 0.9125 and 0.9701 for the model with 5, 6 and 7 factors, respectively, and the adjusted R-Squares keep increasing up to 0.9979. However, models selected by HW2 have the highest adjusted R-Square as 0.6165 in 8-factor model. In their model with 7 variables the R-Square is only 0.8533, which is much less than 0.9701 in our choice. Compared to the final models selected from the censored data, the adjusted R-Squares 0.4985, 0.9642 given by HW1's (E, EG, EH) and our model (4.3), respectively, are smaller than 0.9979 provided by model (4.5). □

[Please place Table 4 here]

Example 6. We use QD method to impute the "corner" data and follow the "mixed" model-selection method to obtain our model and their R-Squares:

$$\text{Intercept, G (.0787), BG (.3112), H (.5105), FH (.7203), DH (.8700), E (.8703),}$$
$$\text{EJ (.9964), DG (.9940), A (.9950), AC (1.0).} \tag{4.6}$$

Since the main effect A and interaction BG have the same contrasts, and the main effect G is induced from the effect BG, which is not s-significant itself, we decide to fix the main effect A as the first term of the model and select additional terms from other factors. This procedure can be done in SAS [11] program PROC REG with "include" option. The final model is given as

Intercept, A (.2326), J (.2572), GJ (.4763), HJ (.7352), JK (.9423), F (.9572),
FC (.9940), K (.9940), GK (0.9996). (4.7)

Since A and K are the most s-significant two main effects, one might include both A and K in the model and proceed the selection. We obtain

Intercept, A (.2326), K (.4120), B (.5021), BE (.7962), J (.9492), B (.9895),
G (.9897), FG (.9965), DG (1.0), AE (1.0), (4.8)

which is slightly better than (4.7). Applying the SME method, we obtain:

Intercept, A (.2326), K (.4120), CK (.5720), JK (.7083), FK (.8717), AK (.9899),
AG (.9964), BA (.9981), HK (.9997), BK (1.0). (4.9)

All these models are comparable but (4.7) and (4.8) are better than (4.9) in the 6th term, where our model has a higher R-Square 0.9423 (or 0.9492) compared to SME's 0.8717. Of course, with more terms, the R-Squares all increase to 1.0. □

4.3.3 Sensitivity Study of the Model Selections

Example 7. To see the impact of different imputations to the selection of models, we impute the "wall" data at the conditional mean of the main-effect model with the parameters estimated from the ML method. The imputed data are given in the original scale as

98.1923, 53.8087, 147.8523, 59.9281, 66.3512, 40.2262, 59.9282,
53.8086, 88.3463, 86.4355, 88.3462, 43.8345. (4.10)

We summarize the model-fitting results in Table 5. The adjusted R-Squares increase to 1.0 in our "mixed" model-selection,

Intercept, E, EG, EH, H, HJ, AH, A, AJ, D, DK, (4.11)

but reach the highest value 0.9866 at 6-variable model in SME. However, HW2's and our procedure suggest the same model up to the 3 variables, which give a good adjusted R-Square 0.9690.

Next, we conduct a cross model-data comparison. Fitting model (4.5) to the imputed data (4.10) from main-effect model gives a comparable result (c.f. the last two columns of Table 5) to that of model (4.11) selected here. We note that the adjusted R-Squares 0.9689, 0.8795, given by HW1's (E, EG, EH) and model (4.3), respectively, are smaller than 1.0 provided by model (4.11). However, surprisingly, fitting of model (4.11) to the data given in Example 5 performs very poorly (c.f the last two columns of Table 4), for example, the highest adjusted R-Square is only 0.6840 for model (4.11) with 6 variables, which is quite low compared to 0.8075 for model (4.5). This is also the case when we fit HW1's model, where the adjusted R-Square is only 0.4985. □

[Please put Table 5 here]

Different imputed data, such as the samples given in Example 5 and 7, indeed suggest different models (2.5) and (2.7), respectively. In next example, we study the sensitivity of the model-selection procedure with respect to the parameter estimates.

Example 8. Since the MLE of model (Intercept, E, EG, EH) exists, we compare the selected models based on data imputed from the following four different estimates under this model assumption: the ML estimates given by HW1, our ML estimates (with initial values provided by LIFEREG), the least squares estimates LSE-1 of (β_0, β_E, β_{EG}, β_{EH}, σ) = (0.984630, −0.4332e-02, 0.1999e-02, −0.1768e-02, 9.5e-04), obtained from fitting the data in Example 6, and the LS-2 estimates, (0.982694, −0.6012e-02, 0.3812e-02, −0.3592e-02, 74.0e-04) from fitting the data in Example 5. We note that the estimate LSE-2 is very different to the first three estimates, especially the estimate of σ is one digit off the others. The first three estimates of regression coefficients β_i's are close, but HW1's ML estimates is larger than ours and they both are larger than LSE-1. The ML estimates of σ from HW1's and ours are close and both are much smaller than the LS estimates. Different parameter estimates gives different imputed data and their models can be different. Table 6 shows four sets of imputed data in the original scales. The imputed data based on LSE-2 is quite different to the others. The final models selected from our procedure are listed in Table 7. The first three models are almost the same except that the order of two terms in the model with HW1's estimates are different to the other two models. The last model with the LSE-2 is very different to the others and the fitting is not good until it includes at least 7 terms in the model. □

[Please place Table 6 and 7 here]

From the study in these examples, we learn that the first three terms E, EG and EH in the "wall" data and the first four terms A, K, E, BE in the "corner" data are the dominating factors and the MLE's exist in both cases. However, in finding the best process recipe, and in doing the model diagnosis, one likes to build a model with a few additional terms, which gives a better prediction for optimization (see next two sections for examples).

5. EXAMPLE 9: SELECTING THE FINAL MODEL FOR THE "WALL" DATA

The MLECHK program indicates that the MLE's exist in the model (4.3) including up to the term J only. LIFEREG gives converged ML estimates of β_0, β_E, β_D, β_{CD}, β_J, σ with log-likelihood -14.482300. We then impute the censored data at the conditional mean [c.f. Eq. (3.2)] and perform the "mixed" model-selection based on these complete samples. The selected model is given in the following sequence:

Intercept, E, EG, EH, F, FK, FB, J, AJ, H, (5.1)

where the MLE exists for all terms in the model. LIFEREG gives an unconverged log-likelihood as -0.24822, but several iterations of Powell's program give a reasonable ML estimates with log-likelihood = 0.0. We then impute the censored data and select the next model:

Intercept, E, EG, EH, D, DH, EJ, DJ, C, CK, (5.2)

where the MLE exists for all terms in the model, too. LIFEREG gives a converged ML estimates for the model up to the term DJ (not C and CK) with log-likelihood =

-0.045274. The ML estimates in the full model are not satisfactory. Since the model did improve the likelihood, we keep searching a possible better model. With the MLE from LIFEREG, we impute the data and select another model:

$$\text{Intercept, E, EG, EH, D, DH, DJ, BD, CD, DG,} \tag{5.3}$$

where the MLE exists up to the term CD (but not DG). LIFEREG gives the converged parameter estimates for the model up to the term BD as

$$\begin{gathered} 0.98405,\ -0.41150\text{e-}02,\ 0.22100\text{e-}02,\ -0.26108\text{e-}02,\ -0.27000\text{e-}03, \\ 0.11253\text{e-}02,\ 0.66286\text{e-}03,\ 0.47180\text{e-}04 \text{ and } \hat{\sigma} = 0.6014\text{e-}04 \end{gathered} \tag{5.4}$$

with the log-likelihood = -1.136512e-09. Running through the Powell's program, we found that the log-likelihoods are all equal to 0.0 in the full models of (5.1), (5.2) and (5.3) up to CD term. In our experience, when Powell's program gives zero log-likelihood, the parameter estimates would not converge even with more iterations of Powell's program itself. Hence, we only compare models with converged ML estimates. Comparing models (4.3) and (5.1) to (5.3), we note that the model changes from E, D, CD, ..., etc. to E, EG, EH, *D, DH,* ..., etc. and the best converged parameter estimates is in model (5.3).

At a particular combination $\underset{\sim}{x}$ of factor-levels of the model (5.3), the mean response is estimated as $\underset{\sim}{x}'\hat{\underset{\sim}{\beta}}$, where $\hat{\underset{\sim}{\beta}}$ is the ML estimate of regression parameter given in (5.4). To find the best combination $\underset{\sim}{x}_{opt}$ of design levels, we conduct a 64-run point search by changing the levels of the factors B, D, E, G, H, J involved in (5.3). The best three combinations of main effects are:

$$\begin{gathered} \text{B}(+1), \text{D}(+1), \text{E}(-1), \text{G}(-1), \text{H}(+1), \text{J}(+1), \text{ predict lifetime } 183.90, \\ \text{B}(-1), \text{D}(+1), \text{E}(-1), \text{G}(-1), \text{H}(+1), \text{J}(+1), \text{ predict lifetime } 180.76, \\ \text{B}(+1), \text{D}(+1), \text{E}(-1), \text{G}(-1), \text{H}(+1), \text{J}(-1), \text{ predict lifetime } 147.85. \end{gathered} \tag{5.5}$$

We see that these combinations are quite similar except the level of B and J might be changed and the levels of E, G and H are the same as the ones suggested by HW1.

However, after fixing the "right" levels in factors B, D and J, we get much higher predicted lifetimes than the prediction 146.18 given in HW1.

Next, we compare our choice (5.5) of factor-levels to E(−1), G(−1), H(+1) from HW1's analysis, and E(−1) from Specht's (1985) application of Taguchi's method. The ML estimates (5.4) are used to obtain 8 and 32 predictions from all possible combinations of the factors (B, D, J) and (B, D, G, H, J) for HW1's and Specht's suggestions, respectively. The minimum, maximum, mean, s.d., c.s and c.k of HW1's and Specht's predictions are computed as (116.71, 183.90, 146.37, 25.04, 0.35, −1.47) and (53.27, 183.90, 94.55, 35.12, 1.04, 0.20), respectively. Our choice (5.5) of the process recipe is the maximum of these predictions and is much larger than their average predictions 146.37 and 94.55. This example shows that different analyses lead to different process recipes and after fixing the "right" level in factors B, D, and J, our prediction is higher than theirs. This conclusion is similar to the result given in HW1, where they suggested the combination E(−1), G(−1) and H(+1) for "wall" data and demonstrated that it is better than Specht's [13] choice E(−1) suggested by Taguchi's method, where the factor G and H are not fixed in the "correct" level.

6. EXAMPLE 10: SELECTING THE FINAL MODEL FOR THE "CORNER" DATA

The MLE of the initial model (4.4) exist up to E (but not EJ). From running Powell's program, we obtain the converged ML estimates of the model parameters as

$$\begin{aligned}&0.99087,\ -0.34965\text{e-}02,\ -0.33184\text{e-}02,\ 0.69832\text{e-}03,\ -0.25884\text{e-}04,\ 0.19904\text{e-}02,\\ &\quad 0.24070\text{e-}02,\ 0.28718\text{e-}02 \text{ and } \hat{\sigma} = 0.70086\text{e-}03 \qquad (6.1)\end{aligned}$$

with the log-likelihood = −7.643244. Since in this case the random imputation suggests a better model, we report its result here. Based on the ML estimates (6.1), the censored data are imputed as:

167.4908, 164.0000, 41.9617, 104.7816, 83.6594, 102.3463, 261.1159,

$$70.2502,\ 365.1984,\ 57.5431,\ 130.7951,\ 9494.7459. \tag{6.2}$$

Applying the "mixed" model-selection procedure, we select the following model:

$$\text{Intercept, A, K, D, DJ, AJ, C, CH, BD, CG,} \tag{6.3}$$

where K is picked up first and E, D, G are induced main effects and the MLE's exist up to the term BD. The MLE's of the parameters in (6.3) (without CG) are

$$\begin{aligned}&0.99047,\ -0.26874\text{e-}02,\ -0.45901\text{e-}02,\ -0.15597\text{e-}02,\ -0.25295\text{e-}02,\ -0.11129\text{e-}02,\\ &\quad -0.61255\text{e-}03,\ -0.69568\text{e-}03,\ 0.13579\text{e-}02 \text{ and } \hat{\sigma} = 0.28489\text{e-}03\end{aligned} \tag{6.4}$$

with the log-likelihood $= -0.376558$ from Powell's program. The censored data are randomly imputed as

$$\begin{aligned}&384.8692,\ 261.0671,\ 39.7295,\ 94.3484,\ 85.9110,\ 103.4602,\ 323.3663,\\ &\quad 61.1780,\ 204.4357,\ 60.8406,\ 135.8842,\ 439.4472.\end{aligned} \tag{6.5}$$

The following model is then selected from the stepwise regression:

$$\begin{aligned}&\text{Intercept, A (.3944), K (.7119), E (.7878), BE (.9488), D (.9926),}\\ &\quad \text{EH (.9997), JK (1.0), GK,}\end{aligned} \tag{6.6}$$

where the MLE's exist up to the term D only and the log-likelihood reported from Powell's program is -5.080860, which is not as good as -0.376558 in (6.3). Hence, we stop the search and treat (6.3) as our final model.

If one takes the conditional mean approach to impute the censored data, the initial model might be selected as

Intercept, A, K, E, BE, G, FG, DG, AH, AK, CK,

which is the same as (6.6) up to the BE interaction and its log-likelihood is not as good as the one for model (6.3).

To find the best combination $\underset{\sim}{x}_{opt}$ of design levels, we conduct a 128-run point search by changing the levels of the factors B, A, C, D, H, J, K involved in (6.3). The best four combinations of the main effects are:

$$B(-1),\ A(-1),\ C(-1),\ D(-1),\ H(+1),\ J(+1),\ K(-1),\ \text{predict } 1.005616, \tag{6.7}$$

$$B(-1),\ A(-1),\ C(+1),\ D(-1),\ H(-1),\ J(+1),\ K(-1),\ \text{predict } 1.004391,$$

$$B(-1),\ A(-1),\ C(-1),\ D(-1),\ H(-1),\ J(+1),\ K(-1),\ \text{predict } 1.004224,$$

$$B(-1),\ A(-1),\ C(+1),\ D(-1),\ H(+1),\ J(+1),\ K(-1),\ \text{predict } 1.002999.$$

These predictions are transformed as infinities in the original scale. We note that only C and H are changing signs in these four combinations.

Next, we compare our choice (6.7) to Specht's [13] conclusion

$$F(+1),\ B(-1),\ A(-1),\ D(-1),\ J(+1),\ K(-1), \tag{6.8}$$

which was suggested from Taguchi's method. Four combinations of factors C and H are given above and their mean, s.d. are computed as 1.004308 and 0.001070, respectively. Our choice (6.7) of the process recipe is the maximum of these predictions.

7. CONCLUSION

In this article, we reanalyze Specht's [13] heat exchanger life data to show that, with a better statistical modeling, the combination of process variables can be determined more accurately. Although there are problems with the maximum likelihood estimation in analyzing such censored data from screening experiments, we are able to compare many models and select a very best one for optimizing the process control. The specification of the initial model and the model selection procedures in Hamada and Wu's [7] algorithm are improved and illustrated with real examples. A sensitivity study of the selected models with the change of imputations and parameter estimates shows the importance of using good models and good estimates in the "imputation-modeling-maximization" procedure. Finally, our model serves as a tool for searching the best combination of process variables. The superiority of the proposed method is demonstrated by comparing our predictions to the existing ones suggested from Taguchi's and Hamada and Wu's procedures.

ACKNOWLEDGMENTS

We thank Dennis Boos for the constructive comments in earlier versions of this manuscript, and Mike Hamada for providing some useful software and Specht's data.

REFERENCES

[1] Box, G. E. P., and Cox, D. R. (1964), "An analysis of Transformations" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 26, 211-252.

[2] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977), "Maximum Likelihood From Incomplete Data Via the EM Algorithm" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 39, 1-38.

[3] Hahn, G. J., Morgan, C. B., and Schmee, J. (1981), "The Analysis of a Fractional Factorial Experiment With Censored Data Using Iterative Least Squares," *Technometrics*, 23, 33-36.

[4] Hamada, M., and Tse, S. K. (1988), "A Note on the Existence of Maximum Likelihood Estimates in Linear Regression Models Using Interval-Censored Data," *Journal of the Royal Statistical Society*, Ser. B, 50, 293-296.

[5] Hamada, M., and Tse, S. K. (1989a), "When Do Estimates Exist From Censored Data in Industrial Experiments?" Research Report 89-03, University of Waterloo, Institute for Improvement in Quality and Productivity.

[6] Hamada, M., and Tse, S. K. (1989b), "MLECHK: A FORTRAN Program for Checking The Existence of Maximum Likelihood Estimates From Censored, Grouped, Ordinal and Binary data from Designed Experiments," Research Report 89-09, University of Waterloo.

[7] Hamada, M., and Wu, C. F. J. (1991), "Analysis of Censored Data From Highly Fractionated Experiments," *Technometrics*, 33, 25-38.

[8] Hamada, M., and Wu, C. F. J. (in press), "Analysis of Designed Experiments With Complex Aliasing," *Journal of Quality Technology*.

[9] Lu, J. C. and Unal C. (1992), "Bayesian Analysis of Censored Data from Highly Fractionated Experiments," working manuscript, the Department of Statistics, North Carolina State University.

[10] Powell, M. J. D. (1964), "An Efficient Method for Finding the Minimum of a Function of Several Variables Without Calculating Derivatives," *Computer Journal*, 7, 155-162.

[11] SAS (version 6.03, 1991), a statistical software package marketed by SAS institute Inc. SAS Campus Drive, Cary, NC 27513.

[12] Silvapulle, M. J. and Burridge, J. (1984), "Existence of Maximum Likelihood Estimates in Regression Models for Grouped and Ungrouped Data," *Journal of the Royal Statistical Society*, Ser. B, 46, 100-106.

[13] Specht, N. (1985), "Heat Exchanger Product Design via Taguchi Methods," in *Third Symposium on Taguchi Methods*, Romulus, MI: American Supplier Institute, Inc., pp. 302-318.

[14] Taguchi, G. (1987), *System of Experimental Design*, White Plains, NY: Unipub/Kraus International.

[15] Wei, G. C. F., and Tanner, M. A. (1990), "Posterior Computations for Censored Regression Data," *Journal of the American Statistical Association*, 85, 829-839.

A list of all tables and figures:

Table 1. Design and Data for the Heat Exchanger Experiment

	Factor										Time to Failure	
Run	*F*	*B*	*A*	*C*	*D*	*E*	*G*	*H*	*J*	*K*	*Wall*	*Corner*
1	1	1	1	1	1	1	1	1	1	1	(93.5,105)	(164,∞)
2	1	1	1	1	1	2	2	2	2	2	(42,56.5)	(164,∞)
3	1	1	2	2	2	1	1	2	2	2	(128,∞)	(0,42)
4	1	2	1	2	2	2	2	1	1	2	(56.5,71)	(93.5,105)
5	1	2	2	1	2	1	2	1	2	1	(56.5,71)	(82, 93.5)
6	1	2	2	2	1	2	1	2	1	1	(0,42)	(93.5,105)
7	2	1	2	2	1	2	2	1	2	1	(56.5,71)	(164,∞)
8	2	1	2	1	2	2	1	1	1	2	(42,56.6)	(56.5,71)
9	2	1	1	2	2	1	2	2	1	1	(82,93.5)	(164,∞)
10	2	2	2	1	1	1	2	2	1	2	(82,93.5)	(56.5,71)
11	2	2	1	2	1	1	1	1	2	2	(82,93.5)	(128,140)
12	2	2	1	1	2	2	1	2	2	1	(42,56.5)	(164,∞)

Table 2. Combinations of Parameter Values
Which Give the Maximum Likelihood in (Intercept, E, EG, EH) Model

β_0	β_E	β_{EG}	β_{EH}	σ		*Count*
0.984647	-0.189939e-02	-0.427636e-02	0.228084e-02	1.32459e-04	(2, ..., 9)*	8
		-0.426636e-02	0.229084e-02	1.32459e-04	(2, ..., 6)	5
		-0.425636e-02	0.230084e-02	1.32459e-04	(2, ..., 5)	4
0.984657	-0.190939e-02	-0.427636e-02	0.228084e-02	1.32459e-04	(2, ..., 9)	8
		-0.426636e-02	0.229084e-02	1.32459e-04	(2, ..., 9)	8
		-0.425636e-02	0.230084e-02	1.32459e-04	(2, ..., 6)	5
		-0.424636e-02	0.231084e-02	1.32459e-04	(2, ..., 5)	4
0.984667	-0.191939e-02	-0.427636e-02	0.228084e-02	1.32459e-04	(2, ..., 8)	7
		-0.426636e-02	0.229084e-02	1.32459e-04	(2, ..., 9)	8
		-0.425636e-02	0.230084e-02	1.32459e-04	(2, ..., 9)	8
		-0.424636e-02	0.231084e-02	1.32459e-04	(2, ..., 6)	5

*: σ values = 1.32459e-04, 1.33459e-04, ..., 1.39459e-04.

Table 3. Rank Table for Selecting Model Terms Based On Censored "Wall" Data

Significant Main Effects	*Induced Main Effects*		*Significant Interactions*	*Main Effects*$	*Constrasts*	*p-Value*+
$E_{(1)}^{\#}$				E	−4.418e-03	0.0001
				B	−1.616e-03	0.3155
			$CD_{(3)}$	K	1.382e-03	0.3977
				C	1.053e-03	0.5246
	C	II*		D	6.131e-04	0.7145
	$D_{(2)}$	II		J	6.131e-04	0.7145
				A	−4.957e-04	0.7664
			$CJ_{(5)}$	H	−4.886e-04	0.7702
	$J_{(4)}$	III		F	2.581e-04	0.8776
				G	2.518e-04	0.8808
			$DH_{(6)}$			
	H	I				
			$AJ_{(7)}$			
	A	I				
			$GJ_{(8)}$			
	G	I				
			$BK_{(10)}$			
	$B_{(9)}$	I				
	K	I				

*: Frequency of the main effects used in the model.
+: Contrasts and p-values computed from fitting regression models to censored wall data one factor at a time.
#: The sequence of the factors entering into the model.
$: Some of the p-values of two-factor interactions are listed in Example 2.

Table 4. Comparison of Model-Selection Procedures
Based on Imputed "Wall" Data from the Quick and Dirty Method

	Hamada and Wu's			Model (4.5)			Model (4.11)	
# Variables	Variables	R^2	Adj–R^2	Variables	R^2	Adj–R^2	R^2	Adj–R^2
1	E	0.3612	0.2973	E	0.3612	0.2973	0.3612	0.2973
2	EG	0.5064	0.3967	B	0.4536	0.3322	0.5064	0.3967
3	EH	0.6353	0.4985	BA	0.6437	0.5100	0.6353	0.4985
4	EJ	0.6950	0.5207	F	0.6490	0.4484	0.6523	0.4536
5	DE	0.7548	0.5505	FA	0.7834	0.6029	0.7304	0.5058
6	FE	0.8089	0.5796	FK	0.9125	0.8075	0.8563	0.6840
7	AE	0.8533	0.5966	BJ	0.9701	0.9178	0.8835	0.6795
8	CE	0.8954	0.6165	FG	0.9884	0.9575	0.9032	0.6452
9	EK	0.9253	0.5892	BC	0.9980	0.9893	0.9033	0.4680
10	BE	0.9481	0.4291	J	0.9998	0.9979	0.9563	0.5193

Table 5. Comparison of Model-Selection Procedures Based on Imputed "Wall" Data from the Main-Effect Model

	Hamada and Wu's			Model (4.11)			Model (4.5)	
# Variables	Variables	R^2	Adj−R^2	Variables	R^2	Adj−R^2	R^2	Adj−R^2
1	E	0.7085	0.6794	E	0.7085	0.6793	0.7085	0.6793
2	EG	0.8594	0.8282	EG	0.8594	0.8281	0.8154	0.7744
3	EH	0.9774	0.9689	EH	0.9774	0.9690	0.8915	0.8508
4	CE	0.9859	0.9778	H	0.9776	0.9647	0.9029	0.8474
5	FE	0.9906	0.9828	HJ	0.9919	0.9851	0.9588	0.9245
6	EK	0.9939	0.9866	AH	0.9964	0.9920	0.9835	0.9636
7	EJ	0.9951	0.9865	A	0.9966	0.9905	0.9920	0.9780
8	DE	0.9963	0.9864	AJ	0.9992	0.9972	0.9928	0.9735
9	AE	0.9969	0.9830	D	0.9997	0.9983	0.9977	0.9871
10	BE	0.9971	0.9681	DK	1.0	1.0	0.9993	0.9920

Table 6. Imputed Data Based on Different Estimates of Model Parameters

Run	*Our MLE*	*HW1's MLE*	*LSE–1*	*LSE–2*
1	94.4615	94.2175	98.1118	98.8972
2	52.0077	52.0861	51.1779	47.8828
3	145.4247	146.1775	146.6097	4243.9542*
4	64.8966	65.1678	62.7998	62.9238
5	65.4584	65.3381	65.9317	62.7072
6	41.8088	41.8573	40.8894	30.2704
7	64.8966	65.1678	62.7998	62.9238
8	50.0230	50.1128	50.1077	47.8270
9	87.2746	87.3286	87.6305	87.3714
10	87.2746	87.3286	87.6305	87.3714
11	92.5578	92.7945	88.2989	87.8329
12	42.1929	42.1444	43.6421	46.9544

*: Estimated mean responses at the third run are 0.993124, 0.993159, 0.992188, 0.996110 for these four estimates, respectively. Since the mean 0.996110 of LSE-2 is so high that the imputaion of the data $(128, \infty)$ becomes large.

Table 7. Selected Models from Data Imputed with Different Estimates

	Our MLE			HW1's MLE			LSE−1			LSE−2		
		R^2	Adj−R^2		R^2	Adj−R^2		R^2	Adj−R^2		R^2	Adj−R^2
1	E	.6717	.6388	E	.6669	.6336	E	.7012	.6714	E	.5308	.4839
2	EG	.8653	.8354	EG	.8629	.8325	EG	.8686	.8394	K	.6158	.5304
3	EH	.9999	.9998	EH	.9999	.9999	EH	.9939	.9917	FK	.8234	.7572
4	K	.9999	.9998	B	.9999	.9999	K	.9942	.9908	J	.8241	.7235
5	AK	.9999	.9999	BD	1.0	.9999	AK	.9984	.9970	CJ	.9417	.8931
6	B	.9999	.9999	K	1.0	.9999	B	.9985	.9966	G	.9417	.8717
7	BD	1.0	1.0	AK	1.0	1.0	BD	.9998	.9993	BG	.9870	.9642
8							J	.9999	.9998	FG	.9951	.9821
9							CK	1.0	1.0	DG	.9986	.9978

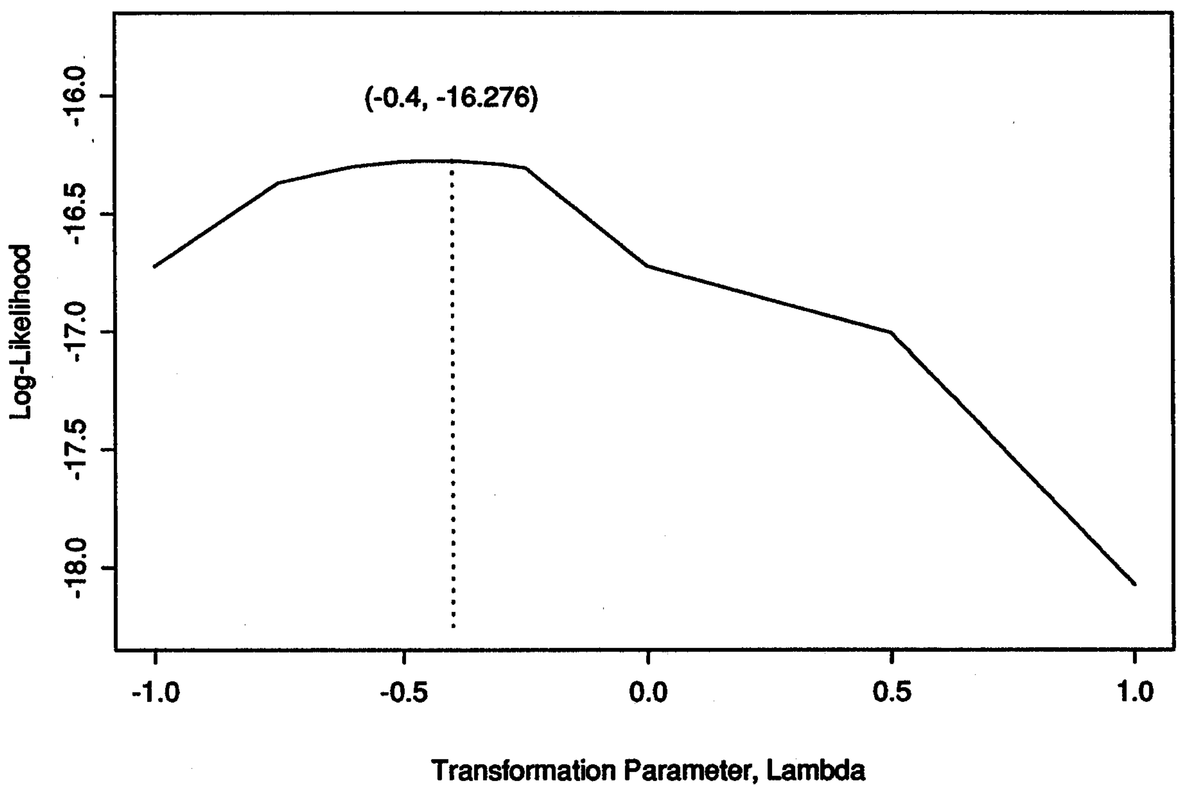

Figure 1, Log-Likelihood of Model (Intercept, A, K)
Against Transformation Parameter Plot