# Battery Life Estimator Manual

# Linear Modeling and Simulation

Jon Christophersen
Ira Bloom
Ed Thomas
Vince Battaglia

August 2009

**INL**
Idaho National Laboratory

# Battery Life Estimator Manual

# Linear Modeling and Simulation

Jon Christophersen
Ira Bloom[1]
Ed Thomas[2]
Vince Battaglia[3]

[1]ANL
[2]SNL
[3]LBNL

**August 2009**

**Idaho National Laboratory**
**Energy Storage and Transportation Systems Department**
**Idaho Falls, Idaho 83415**

**http://www.inl.gov**

# Battery Life Estimator Manual

## Linear Modeling and Simulation

August 2009

Idaho National Laboratory
Idaho Falls, ID  83415
Operated by Battelle Energy Alliance, LLC

Argonne National Laboratory
Argonne, IL  60439
Operated by the University of Chicago

Sandia National Laboratories
Albuquerque, NM  87185
Operated by Lockheed Martin Corporation

Lawrence Berkeley National Laboratory
Berkeley, CA 94720
Operated by University of California

*FreedomCAR & Vehicle Technologies Program*

# CONTENTS

# FIGURES

# TABLES

# GLOSSARY OF TERMS

*Beginning of Life (BOL)* – The point in time at which life testing begins. A distinction is made in this manual between the performance of a battery at this point and its initial performance, because some degradation may take place before the start of life testing. Analysis of the effects of life testing is based on changes from the BOL performance.

*Calendar Life* – The time required to reach end of life at the reference temperature at open-circuit (corresponding to key-off/standby conditions in the vehicle).

*Cycle Life* – The number of consecutive cycles consisting of a charge neutral combination of discharge and charge pulses centered on a given state-of-charge required to reach end of life at the reference temperature.

*Degradation Model* – An empirical or chemistry/physics-based model that describes the expected degradation of a battery experiencing typical stress conditions.

*Depth of Discharge (DOD)* – The percentage of a device's rated capacity removed by discharge relative to a fully charged condition, normally referenced to a constant current discharge at the $C_1/1$ rate. The capacity to be used is established (fixed) at the beginning of testing, %.

*End of Life (EOL)* – A condition reached when the device under test is no longer capable of meeting the applicable USABC goals. This is normally determined from RPT results, and it may not coincide exactly with the ability to perform the life test profile (especially if cycling is done at elevated temperatures.) The number of test profiles executed at end of test is not necessarily equal to the cycle life per the USABC goals.

*End of Test (EOT)* – The point in time where life testing is halted, either because criteria specified in the test plan are reached, or because it is not possible to continue testing.

*Error Model* – A model that accounts for the difference between the measured and expected performance. The error model combines the effects of both measurement error and manufacturing variability.

*Reference Performance Test (RPT)* – A periodic assessment of battery degradation during life testing. A reference performance test will typically yield capacity fade, power fade, and impedance rise as a function of test time.

*State of Charge (SOC)* – The available capacity in a battery expressed as a percentage of actual capacity. This is normally referenced to a constant current discharge at the $C_1/1$ rate. For this manual, it may also be determined by a voltage obtained via a relationship of capacity to voltage established at beginning of life. SOC = (100 – DOD) if the rated capacity is equal to the actual capacity, %.

*Stress Conditions* – The parameters that are used to accelerate aging of a battery technology, such as temperature, state-of-charge, throughput, and pulse power. These are the explanatory variables in the degradation model.

# ACRONYMS

| | |
|---|---|
| BLE | Battery Life Estimator |
| BOL | beginning of life |
| CDF | cumulative distribution function |
| DOD | depth of discharge |
| EOL | end of life |
| EOT | end of test |
| LCL | lower confidence limit |
| LOF | lack of fit |
| MAV | median absolute value |
| MCS | Monte Carlo simulation |
| RPT | reference performance test |
| SOC | state of charge |
| SS | sum of squares |
| UCL | upper confidence limit |
| USABC | United States Advanced Battery Consortium |

# Battery Life Estimator Manual

## 1   INTRODUCTION

The purpose of this Battery Life Estimator (BLE) Manual is to assist developers in their efforts to determine the life capability of advanced battery technologies for automotive applications.   Testing requirements and procedures have been previously defined in manuals published under the United States Advanced Battery Consortium (USABC) in References 1 through 5.   This manual describes a standardized method of determining calendar life with a high degree of statistical confidence based on models and degradation data acquired from typical battery testing.

A software package ("BatteryLife.exe") has also been developed to estimate calendar life based on the methodology described herein.   To purchase the software, users must first download and sign the license agreement at http://www.anl.gov/techtransfer/ Software_Shop/TLVT/TLVT.html.   USABC developers can acquire the software at no cost.   The degradation model presented in this manual has been included as the default for quick implementation to a set of data.   However, the software can also accommodate any degradation model that is applicable to a particular chemistry as long as it is linear or can be linearized by appropriate mathematical transformations.

This manual is organized in two main sections.   Section 2 describes the default statistical models implemented in the software as well as the methods for estimating model parameters and cell life from experimental data.   This section also presents a methodology for assessing the uncertainty of the estimated cell life using Monte Carlo simulations.    Section 3 is a user's guide for the software tool ("BatteryLife.exe") and provides details on data formatting, menu navigation, and data processing.   Appendix A is an extended discussion of parameter estimation using robust linear regression.   An applied example of this life estimation methodology using data from a set of test cells is provided in Reference 6.

## 2   METHODOLOGY FOR ESTIMATING CALENDAR-LIFE

This section describes a methodology for estimating the average calendar life of various cell technologies and assessing their readiness for transition to production. Consequently, the emphasis is placed on predicting the capability of *typical* (i.e., representative) cells to meet the USABC target of a 15-year calendar life.   A two-part model can be constructed from the experimental test data.    The first part is the degradation model that represents the average cell performance as a function of aging over a range of stress conditions.  The second part is the error model that represents the deviation of the cell behavior relative to the average performance.   The degradation model provides a basis for the estimation of average cell life, and the error model provides a basis for assessing the accuracy of the degradation model.

At this stage of life testing, there will generally be incomplete knowledge of the specific degradation mechanisms or the source of the deviations between average

performance and actual measured performance. Thus, relatively simple empirical models with few parameters should be used for life estimation, though physics-based models could be used as well. Simple forms of a degradation model and error model that have been successfully implemented for a variety of technologies are illustrated herein. The methodology for estimating model parameters, assessing model accuracy, and estimating mean cell life with associated uncertainty are also described.

## 2.1  Generalized Model

The use of accelerated degradation testing to verify life capability requires the selection of performance measures that accurately reflect battery state of health. An example performance measure is relative resistance (*i.e.*, the cell resistance at time *t* divided by the resistance at beginning of life, $t = 0$). The generalized model must relate the measured cell performance at any given time to a combination of the stress factor effects. For example, in the case of calendar-life experiments with a single stress factor of temperature, the acquired data can be represented generically by the model shown in Equation (1), where $Y_i(T;t)$ represents the measured performance of the $i^{th}$ cell after being subjected to aging for time *t* at temperature *T* (Reference 6). The average cell performance is represented by a degradation model, $\mu(T;t)$, which is described in Section 2.2. The combined effects that are related to the unique behavior of the $i^{th}$ individual cell and measurement error are represented by an error model $\gamma_i(T;t)$, which is described in Section 2.3.

$$Y_i(T;t) = \mu(T;t) + \gamma_i(T;t) \tag{1}$$

## 2.2  Degradation Model

The degradation model can be empirical, chemistry/physics-based, or some combination of both. A wide variety of model forms are possible. The specific form of the model will necessarily depend on the particular technology and set of stress factors.

In the example case of a single stress factor (temperature), a simple but useful form for the degradation model is given by Equation (2), where $\beta_0, \beta_1$, and $\rho$ represent the model parameters (Reference 6).

$$\mu(T;t) = 1 + \exp\left\{\beta_0 + \beta_1 \cdot \frac{1}{T}\right\} \cdot t^\rho \tag{2}$$

Note that $\mu(T;t) = 1$ for $t = 0$ and then *increases* in value as the cell ages. Various normalized responses (including relative resistance) are consistent with these conditions. When the natural response *decreases* to zero as a function of cell age, $\mu(T;t)$ can be considered as a model for the inverse of the natural response. Examples of a naturally decreasing response include relative power and relative capacity. In such cases, $\mu(T;t)$ can be considered as a model for inverse relative power or inverse relative capacity.

To estimate the parameters associated with the degradation model, it is useful to re-express the model in a linear form with a log transformation as shown in Equation (3). Once the model has been linearized, robust linear regression can be used to estimate the model parameters (see Section 2.4).

$$\log(\mu(T;t)-1) = \beta_0 + \beta_1 \cdot \frac{1}{T} + \rho \cdot \log(t) \tag{3}$$

## 2.3  Error Model

The error model accounts for the difference between the measured performance and expected performance. The difference is a combination of effects due to measurement error as well as the intrinsic difference in performance between cells. Two different approaches for determining the measurement error are discussed in this section. One method estimates the error from the measured data, and the other method independently determines the error based on calibration and accuracy checks of the test equipment. The software tool ("BatteryLife.exe") provides both options to the user for life prediction.

### 2.3.1  Estimated Error Model

In the example case of a single stress factor (temperature), a useful form for the estimated error model (Reference 6) is given by Equation (4), where $\delta_i$ represents a random, cell-specific, proportional effect with variance $\sigma_\delta^2$, and $\pi_i(t)$ represents the effects of measurement error on $Y_i(T;t)$.

$$\gamma_i(T;t) = \delta_i \cdot (\mu(T;t)-1) + \pi_i(t) \tag{4}$$

Using relative resistance as the performance measure, the expression for $Y_i(T;t)$ is as shown in Equation (5), where $R_{true}(i,t)$ is the unknown (but true) value of the resistance of the $i^{th}$ cell at time $t$, and $\varepsilon_i(t)$ is the specific unknown error associated with that measurement.

$$Y_i(T;t) = \frac{R_{true}(i,t) + \varepsilon_i(t)}{R_{true}(i,0) + \varepsilon_i(0)} \tag{5}$$

The resulting error model due to measurement effects is shown in Equation (6).

$$\pi_i(t) = \frac{R_{true}(i,t) + \varepsilon_i(t)}{R_{true}(i,0) + \varepsilon_i(0)} - \frac{R_{true}(i,t)}{R_{true}(i,0)} \tag{6}$$

It is assumed that the measurement errors are independent with a relative standard deviation of $\alpha$ (i.e. $\sigma_\varepsilon = \alpha \cdot R_{true}(i,0)$). With this and other assumptions (Reference 6), the variance of $\pi_i(t)$, given by $\sigma_\pi^2$, can be approximated by $2 \cdot \alpha^2$.

3

Assuming that the mean values of $\varepsilon_i(t)$ and $\delta_i$ are zero, then within a given group of cells that have experienced the same stresses and aging time, the mean and variance of $Y_i(T;t)$ can be expressed as shown in Equations (7) and (8). Robust linear regression (see Section 2.4) is used to estimate the variance model parameters ($\sigma_\delta^2$ and $\sigma_\pi^2$).

$$Mean(Y_i(T;t)) = \mu(T;t) \tag{7}$$

$$Var(Y_i(T;t)) = Var(\gamma_i(T;t)) \approx \sigma_\delta^2 \cdot (\mu(T;t)-1)^2 + \sigma_\pi^2 \tag{8}$$

Thus, this model of the variance within a treatment group and Reference Performance Test (RPT) (References 1-5), implies that the expected variability in cell performance increases as the expected level of degradation increases.

### 2.3.2  Independent Assessment of Measurement Error

Alternatively, the magnitude of measurement error can be estimated directly using the uncertainty methodology developed at the Idaho National Laboratory (References 7 and 8). First, the effect of measurement error can be minimized with test equipment calibration and verification. Calibration can be performed using the manufacturer's recommended procedures. Verification consists of independent measurements of test channel voltage and current outputs at various levels within the channel full scale operating range. The total equipment and channel error can then be determined by the measured data and the uncertainties of the independent measurement equipment (*i.e.*, a digital voltmeter and shunt for current measurements). If the results from this analysis show poor accuracy or repeatability, the test equipment should be calibrated and verified again until the results are less than or equal to the claimed values of the manufacturer (*e.g.*, 0.02% of full scale repeatability).

These data are also useful in determining the uncertainty range of the performance parameters of interest for the life prediction model. Each performance parameter (*e.g.*, resistance, power, capacity, and energy) is a function of voltage and current measurements (temperature uncertainty is treated elsewhere). The uncertainty expression associated with that performance parameter can be determined based on the accuracy and precision of the voltage and current measurements as determined during the initial calibration or in-test calibration checks, and low-order Taylor Series approximations of the performance parameter with respect to the independent voltage and current measurements. For example, the uncertainty expression for resistance as defined in the USABC Manuals (References 1-5) is given by Equation (9), where $V_{FS}$ and $I_{FS}$ are the test channel's full scale voltage and current range, respectively; %errV$_{CAL}$ and %errI$_{CAL}$ are the calibration errors due to the independent digital voltmeter and shunt used to measure the accuracy during the calibration check; and %errV$_{STD}$ and %errI$_{STD}$ are the standard deviations determined experimentally from the accuracy measurements.

$$\%R_S = \left[ 2\left( \frac{\%errV_{STD}}{V(t_a)-V(t_b)}V_{FS} \right)^2 + 2\left( \frac{\%errI_{STD}}{I(t_a)-I(t_b)}I_{FS} \right)^2 + (\%errV_{CAL})^2 + (\%errI_{CAL})^2 \right]^{\frac{1}{2}} \quad (9)$$

## 2.4  Robust Linear Regression

The parameters associated with both the degradation and error models are estimated with a robust linear regression procedure because it has reduced sensitivity to anomalous data (*i.e.*, outliers). Consequently, the parameter estimates are not greatly affected by the outliers. Robust regression procedures also are valuable when the error variance is not constant across the experimental space, as is the case for the assumed error model in Section 2.3. The particular procedure implemented in this manual includes three iterations of weighted least-squares regression (Reference 9). For the first iteration, ordinary least-squares regression is used (*i.e.*, the relative weights are identical). For subsequent iterations, the weights are based on Tukey's biweight function (Reference 10). More details concerning parameter estimation are provided in Appendix A.

## 2.5  Life Prediction

The fitted degradation model can be used to estimate the mean lifetime of the cell at a specified temperature for a given end-of-life criterion. Given the degradation model provided in Equation (2), and an end-of-life criterion defined to be a 30% increase in degradation of the performance measure (*i.e.*, $\mu(T;t)$ becomes 1.3 at a target temperature of $T_0$), the resulting estimated lifetime, ($\hat{t}_{EOL}$), is shown in Equation (10).

$$\hat{t}_{EOL} = \exp\left\{ \frac{\log(0.3) - \left\{ \hat{\beta}_0 + \hat{\beta}_1 \cdot \frac{1}{T_0} \right\}}{\hat{\rho}} \right\} \quad (10)$$

## 2.6  Monte Carlo Simulations

Monte Carlo simulations based on the fitted degradation and error models, in conjunction with a variant of the parametric bootstrap procedure (Reference 11), are used to assess the uncertainty of the cell life and associated model parameters. Simulation results provide a basis for assessing the quality of the model based on "lack-of-fit" statistics (Section 2.7). Assuming that the forms of the degradation and error models are accurate, the simulations can then be used to assess the uncertainty of the mean cell life as well as the model parameters estimated from experimental data.

Using a performance measure of relative resistance as an illustration, the overall model in Equation (1) can be expanded to the form shown in Equation (11), where *j* represents the stress condition and *ij* represents the $i^{th}$ cell within the $j^{th}$ stress condition.

$Y_{ij}(t)$ represents the *measured* relative resistance of the $ij^{th}$ cell at time $t$ and $\mu\left(X_j;t\right)$ represents the *expected* relative resistance for cells under the $j^{th}$ stress condition at time $t$. $\delta_{ij}$ represents the random proportional effect of the $ij^{th}$ cell, and $\pi_{ij}(t)$ represents the effect of the random measurement errors on relative resistance associated with the $ij^{th}$ cell at the initial measurement and at time $t$. The last term can be notionally partitioned into two terms: $\pi_{ij}(t) \cong \lambda_{ij}(0) + \lambda_{ij}(t)$, where $\lambda_{ij}$ represents the effect of the individual measurement errors on relative resistance at beginning of life and at time $t$. For these simulations, the random effects, $\delta_{ij}, \lambda_{ij}(0),$ and $\lambda_{ij}(t)$, are assumed to be independent and normally distributed each with a mean of zero and variance of $\sigma_\delta^2, \alpha^2,$ and $\alpha^2$, respectively.

$$Y_{ij}(t) = \mu\left(X_j;t\right) + \delta_{ij} \cdot \left(\mu\left(X_j;t\right) - 1\right) + \pi_{ij}(t) \tag{11}$$

The general approach is to repeatedly simulate the experiment while matching the test duration, RPT frequency, experimental conditions, and number of cells per experimental condition of the actual experiment. For each independent simulation trial (representing a single realization of the complete experiment), different random realizations of cell-to-cell effects and measurement errors are added to the assumed truth provided by the degradation model that was fitted to the actual experimental data. First, the number of stress conditions that were used ($J$), the number of cells tested per condition $\{n_j : j = 1:J\}$, and the times at which the cells were measured $\{t_k : k = 1:K\}$ are identified. Next, the degradation model for each combination of stress condition and measurement time can be computed with $\{\mu\left(X_j;t_k\right):(j=1:J)\times(k=1:K)\}$. Finally, using this setup, a number of independent trials are completed as follows:

1. Simulate $\{\delta_{ij} : (i = 1:n_j)\text{ with }(j=1:J)\}$, where the $\delta_{ij}$ are sampled independently from a normal distribution with mean zero and standard deviation, $\sigma_\delta$.

2. Simulate $\{\lambda_{ij}(0):(i=1:n_j)\text{ with }(j=1:J)\}$, where the $\lambda_{ij}(0)$ are sampled independently from a normal distribution with mean zero and standard deviation $\alpha$.

3. Simulate $\{\lambda_{ij}(t_k):(i=1:n_j)\text{ with }(j=1:J)\text{ and }(k=1:K)\}$, where the $\lambda_{ij}(t_k)$ are sampled independently from a normal distribution with mean zero and standard deviation $\alpha$.

4. Combine the constituent effects from Steps 1 to 3 to form the simulated data:
$Y_{ij}(t_k) = \mu\left(X_j;t_k\right) + \delta_{ij} \cdot \left(\mu\left(X_j;t_k\right) - 1\right) + \lambda_{ij}(0) + \lambda_{ij}(t_k)$
   a. Ensure that $Y_{ij}(t_k) > 1$

5. Model the collective set of simulated resistance data for the current trial:
   a. Estimate model parameters (degradation and error)
   b. Estimate average cell life
   c. Compute the lack of fit sum of squares ($SS_{LOF}$) (Section 2.7)

The summary statistics (*e.g.*, standard deviations and order statistics) of model parameters, estimated cell life, and $SS_{LOF}$ across trials can then be computed. The standard deviations of the model parameters and estimated cell life are referred to as bootstrap standard errors.

## 2.7  Lack-of-fit Statistic

It is also important to assess how well the degradation model fits the experimental data (*i.e.*, the level of performance variation observed for cells aged under a common stress condition). Inaccuracies in the degradation model are detected by the lack-of-fit statistic shown in Equation (12), where $J$ is the number of stress conditions, $K$ is the number of RPT's (the beginning of life RPT is denoted as RPT0), $\bar{Y}_{j \cdot t}$ is the average performance measure (*e.g.*, relative resistance) of the $j^{th}$ stress group at RPT$K$ corresponding to some time $t$ (consisting of $n_{jt}$ cells), $\hat{\mu}_{jt}$ is the fitted degradation model for the $j^{th}$ stress group at RPT$K$, and $\hat{\sigma}_{jt}^2$ is the fitted error model for the $j^{th}$ stress group at RPT$K$, as shown in Equation (13). Note that Equation (12) is normalized by the product of the number of stress conditions and RPTs ($J \cdot K$) to enable a comparison across different experiments.

$$SS_{LOF} = \frac{1}{J \cdot K} \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{n_{jt_k}}{\hat{\sigma}_{jt_k}^2} \cdot \left( \bar{Y}_{j \cdot t_k} - \hat{\mu}_{jt_k} \right)^2 \tag{12}$$

$$\hat{\sigma}_{jt}^2 = \hat{\sigma}_{\delta}^2 \cdot \left( \hat{\mu}_{jt} - 1 \right)^2 + \hat{\sigma}_{\pi}^2 \tag{13}$$

Monte Carlo simulations (Section 2.6) based on the developed degradation and error models are used to assess the lack-of-fit statistic. The value of $SS_{LOF}$ based on the original data is compared with the empirical distribution of the $SS_{LOF}$ values obtained via the simulation trials. An unusually large value for the lack-of-fit statistic (*e.g.*, greater than the 95th percentile of the simulated $SS_{LOF}$ values) is indicative of model inaccuracy.

## 2.8  Application to Calendar Life Data

The recommended pathway for demonstrating adequate calendar life of a battery technology given a set of experimental data from typical USABC testing (References 1-5) is as follows:

1. A deterministic degradation model (Section 2.2) is developed to reflect the average (*i.e.*, typical) cell degradation over time as a function of various stress factors such as temperature and state-of-charge. This model must be accurate over the anticipated range of conditions the cells will experience. It is also assumed that the cell technology is sufficiently advanced such that, given a standard reference (*e.g.*, 30°C), the model will predict a life capability exceeding the target requirement with some significant margin (*i.e.*, the lower confidence bound for predicted life must exceed the 15 year goal).

2. An accurate error model (Section 2.3) is developed to reflect the cell-to-cell variability in observed degradation about the average behavior. This model accounts for variability due to measurement error as well as the intrinsic differences between cells.

3. The degradation and error models from Steps (1) and (2) are used as the basis for conducting Monte Carlo simulations (Section 2.6) to assess the lack-of-fit statistic (Section 2.7).

4. If there is no evidence for lack-of-fit, the average cell life is estimated at the lower confidence limit via the fitted degradation model.

5. Adequate calendar life of a battery technology is demonstrated if the lower confidence limit bound of the estimated cell life exceeds the requirement.

# 3 BATTERY LIFE ESTIMATION SOFTWARE

The purpose of the BLE software package ([http://www.anl.gov/techtransfer/Software_Shop/TLVT/TLVT.html](http://www.anl.gov/techtransfer/Software_Shop/TLVT/TLVT.html)) is to facilitate the analysis of simulated and experimental test data. The software captures all of the mathematical analysis and life-simulation tools described in this manual, and includes the default model described in Section 2. This section provides a description of the software contents and a user's guide to running the application.

## 3.1 System Requirements

The minimum system requirements for the BLE software package are a PC computer with a Pentium 4 processor with at least 1 GB of memory and VGA graphics. There should also be a minimum of 3 MB of free space on the hard disk.

The software has been extensively tested on Windows XP, but the minimum operating system must be at least Windows 2000 (it will not work on Windows 98 or earlier). The Microsoft .NET framework versions 1.1, 2.0, 3.0, and 3.5 must be installed for the software to operate. These framework versions can be downloaded from the following websites:

- Version 1.1:
  [http://www.microsoft.com/downloads/details.aspx?familyid=262D25E3-F589-4842-8157-034D1E7CF3A3&displaylang=en](http://www.microsoft.com/downloads/details.aspx?familyid=262D25E3-F589-4842-8157-034D1E7CF3A3&displaylang=en)

- Version 2.0:
  [http://www.microsoft.com/downloads/details.aspx?FamilyID=0856EACB-4362-4B0D-8EDD-AAB15C5E04F5&displaylang=en](http://www.microsoft.com/downloads/details.aspx?FamilyID=0856EACB-4362-4B0D-8EDD-AAB15C5E04F5&displaylang=en)

- Version 3.0:
  [http://www.microsoft.com/downloads/details.aspx?FamilyID=10CC340B-F857-4A14-83F5-25634C3BF043&displaylang=en](http://www.microsoft.com/downloads/details.aspx?FamilyID=10CC340B-F857-4A14-83F5-25634C3BF043&displaylang=en)

- Version 3.5:
  [http://www.microsoft.com/downloads/details.aspx?FamilyID=333325fd-ae52-4e35-b531-508d977d32a6&DisplayLang=en](http://www.microsoft.com/downloads/details.aspx?FamilyID=333325fd-ae52-4e35-b531-508d977d32a6&DisplayLang=en)

## 3.2 Battery Life User Interface

The BLE software application is designed to be very user-friendly. All functions are accessible through menus and data-entry forms. When the application starts, the blank desktop, shown in Figure 1 will appear. The main menu consists of five choices: File, Edit, Run, View and Help. A detailed description of these menu choices are provided in Section 3.5.
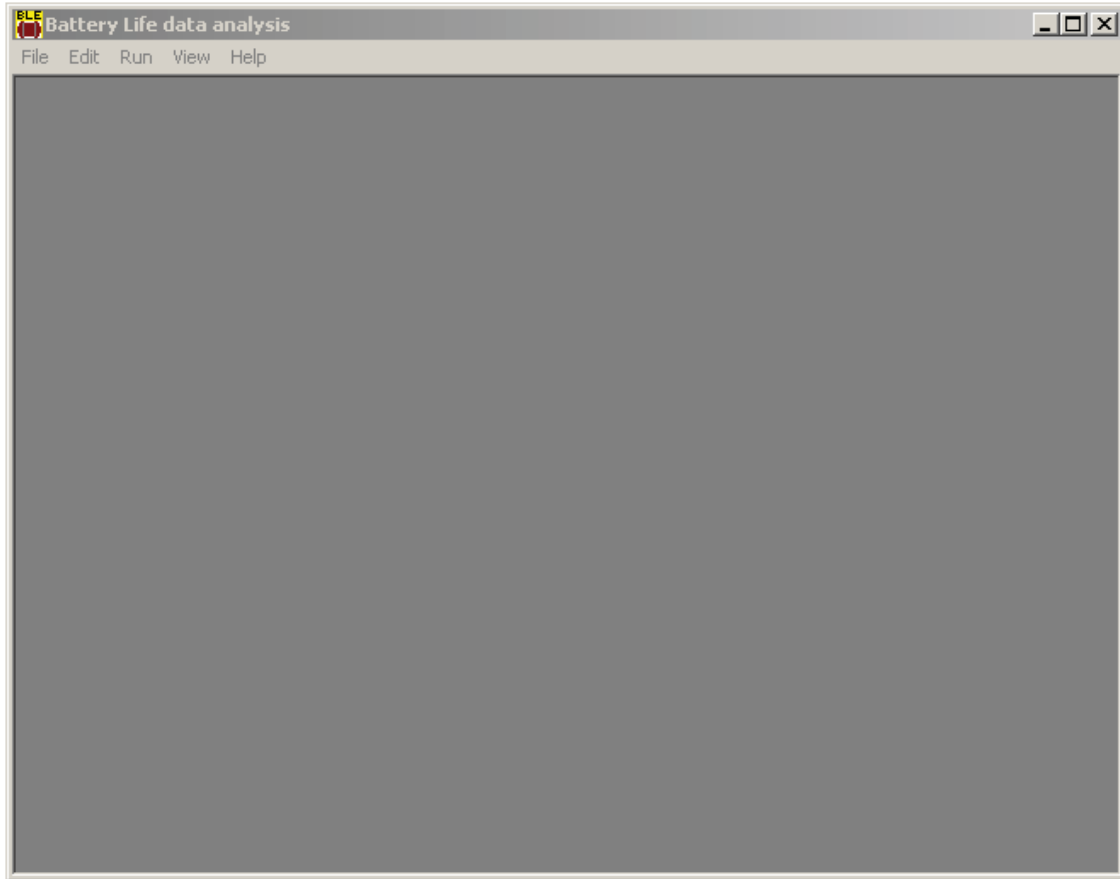
Figure 1.  BatteryLife.exe desktop

## **3.3  Equation Requirements**

The data to be fit or modeled must be time-based, with time as a dependent variable.  Additionally, the modeling equation must be linearized without containing exp(*const*) or ln(*const*) terms, where *const* is a constant.  The resulting transform equations must each consist of only one independent variable.  Consequently, the number of transform equations for the linearized form is based on the number of independent variables in the model equation.

For example, the expression $R = A + B/T + Ct^z$ with $A$, $B$, and $C$ as the dependent variables has the linear transform equations of $R$, $1/T$, and $t^z$.  A more complex expression such as $R=A\exp(B/T)t^z$ can be linearized to $\ln R=\ln A + B/T + z\ln(t)$, but it is not in the proper form because of the $\ln A$ term (log of a constant).  The original expression should be rewritten as $R=\exp(A' + B/T)t^z$, then linearized with $\ln R=A' + B/T + z\ln t$ with transform equations of $\ln R$, $1/T$, and $\ln t$.

# 3.4  Data Formatting Requirements

It is assumed that the data to be fitted are arranged in columns (*e.g.*, in EXCEL or plain text) with a comma as the delimiter in plain-text files.  A header row at the top of each column must be included as well.  The columns do not have to be sequential.  It is further assumed that all data to be fit or modeled are time-based and that time is included as one of the columns in the discrete variable input file.

The data to be fit should be normalized to the $t = 0$ value for each dependent variable of the degradation model (Section 2.2).  However, the beginning of life value (*i.e.*, at $t = 0$) should not be included in the BLE software data file.  If it is not removed, the program will filter these values out as well as any dependent variable values that are less than zero.

Table 3.1 shows part of a sample input file arranged in columns with time, temperature, and relative resistance.  Note that if there are missing measurements, as shown in the second row of data in Table 3.1, they must be removed before starting the data analysis and simulation program.  If they are not, an error message will be displayed, stating that there are no data available.

Table 3.1. Example of data containing a missing measurement.

| Time, yr | Temperature, K | Relative Resistance |
|---|---|---|
| 0.1 | 298 | 1.02 |
| 0.1 | 298 | |
| 0.1 | 298 | 1.01 |
| 0.1 | 308 | 1.03 |

# 3.5  Software Navigation

In the accompanying description, the "▶" symbols indicate a sequence of mouse clicks.

## 3.5.1  File Menu

The File menu is shown in Figure 2 and consists of the following choices: New, Open, Save, Save As, Preview Report, Print Report and Exit.
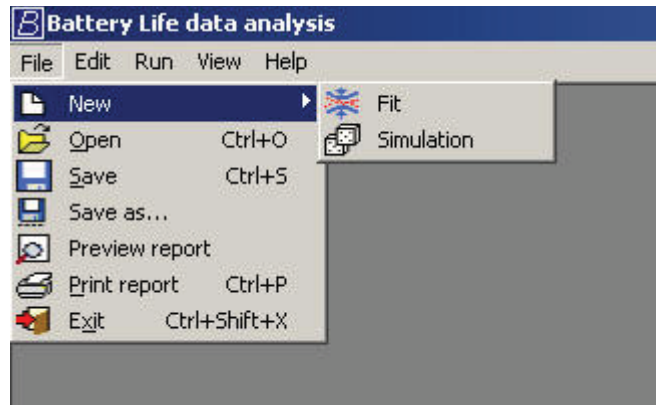


Figure 2.  File menu

File ▶ New ▶ Fit.  This option activates the Fit Wizard (Section 3.6.3) and allows the user to start a new fit based on a modeling equation and an existing set of data (*i.e.*, from a file).

File ▶ New ▶ Simulation.  This option activates the Simulation Wizard (Section 3.6.4) and allows the user to start a new Monte Carlo simulation based on a modeling equation.

File ▶ Open (Ctrl+O).  This option opens an existing *.Life file.  The files contain the information needed to carry out the fit or simulation or a combination of the two.  If the application finds that there is an error in the file, such as it was created by an earlier version of the software or is corrupt, the application will display an error. (Section 3.4)

File ▶ Save (Ctrl+S).  This option saves the existing fit / simulation information to a fit for later use.  The file is saved to the current filename or to a file called "default.Life."

File ▶ Save As.  This option saves the current fit / simulation information to a fit for later use.  The file is saved to a user-defined filename.

File ▶ Preview Report.  This option lets the user see the report on the desktop before committing it to paper (Section 3.7.3).

File ▶ Print Report (Ctrl+P).  This option lets the user commit the fit / simulation calculations to paper.

File ▶ Exit (Ctrl+Shift+X).  This option ends the current session.

### 3.5.2  Edit Menu

The Edit Menu is shown in Figure 3 and consists of the following choices: Equations, Life confidence limit, and Copy view to clipboard.
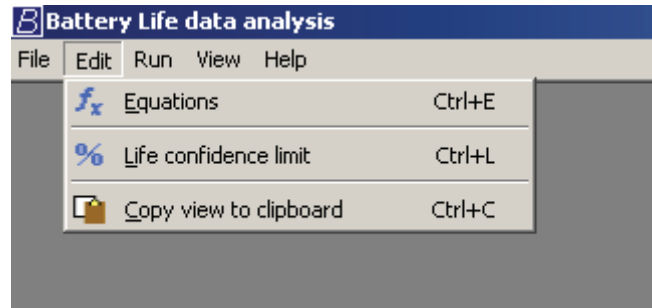


Figure 3.  Edit menu

Edit ► Equations (Ctrl+E).  This option starts the equation editor to alter the stored equations that are used in the current fit and/or simulation (Section 3.6.3.3).

Edit ► Life confidence limit (Ctrl+L).  This option changes the lower and/or upper confidence limits that are used for the life projections (Section 3.6.2.4).

Edit ► Copy view to clipboard (Ctrl+C).  This option copies the contents of the current (foremost) view to the Windows clipboard for pasting into other clipboard-aware applications, such as EXCEL or Word.  It is copied as a windows metafile if the view contains a plot, or as HTML if the view contains text and/or data tables.  Data table displays are copied exactly as they appear on the screen.

### 3.5.3  Run Menu

The Run Menu is shown in Figure 4 and consists of the following choices: Fit and Simulation.
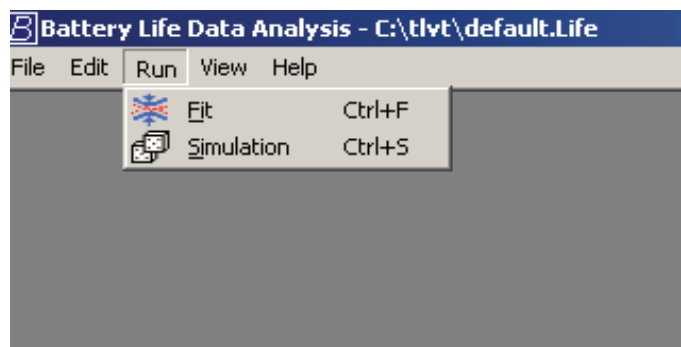


Figure 4.  Run menu

Run ► Fit (Ctrl+F).  This option re-runs the present fit.

Run ► Simulation (Ctrl+S).  This option re-runs the present simulation or creates a simulation based on the present fit.

### 3.5.4  View Menu

The View menu is shown in Figure 5 and consists of the following choices: Results of Models, Plot..., and Data....
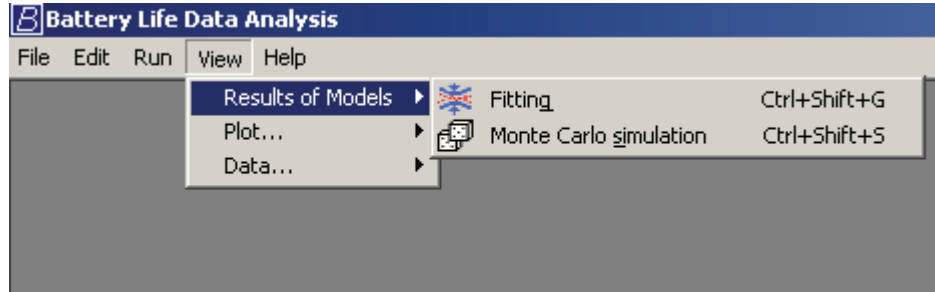


Figure 5.  View menu and model results sub-menu

### 3.5.4.1  *Results of Models*

The sub-menu for model results is also shown in Figure 5.

View ► Results of Models ► Fitting (Ctrl+Shift+G).   This option displays the results of the fitting calculations, and includes the model equation, the names and variables used, the values of the fitting parameters, the estimate of errors, and lack of fit information (Section 3.7.1).

View ► Results of Models ► Monte Carlo simulation (Ctrl+Shift+S).  This option displays the results of the Monte Carlo simulation (Section 3.7.2).

### 3.5.4.2  *Plot*

The plot sub-menu is shown in Figure 6 and has five choices:  Fitting Results, Var(Y) vs. ($\hat{\mu}$-1)², ($\overline{Y}$ - $\hat{\mu}$) vs. $\hat{\mu}$, $\hat{\mu}$ vs. $\overline{Y}$, and Monte Carlo simulation (Section 3.7.2.2).
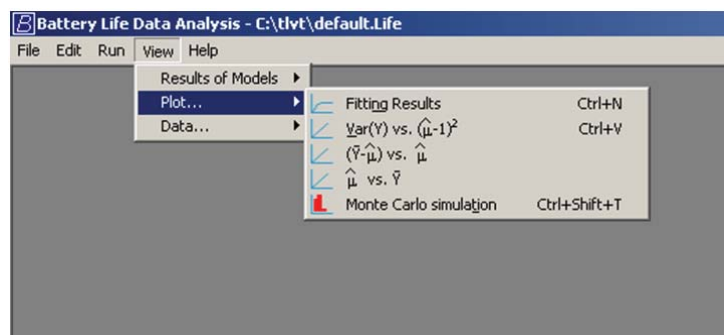


Figure 6. View menu and plot sub-menu

View ► Plot... ► Fitting Results (Ctrl+Shift+N).  This option produces a plot of the experimental data and the fit.

View ► Plot... ► Var(Y) vs. ($\hat{\mu}$ -1)² (Ctrl+Shift+V).  This option produces a plot of the error model results.

View ► Plot... ► ($\overline{Y}$ - $\hat{\mu}$) vs. $\hat{\mu}$.  This option plots the difference between average and predicted performance results (*e.g.*, of relative resistance) versus the predicted performance.

View ► Plot... ► $\hat{\mu}$ vs. $\overline{Y}$.  This option produces a plot of the predicted performance behavior (*e.g.*, predicted relative resistance) versus the average measured performance observed during testing.  A 45° line is also included in this plot to indicate where the data should lie for a perfect fit.  Data that lie well off the 45° line should be evaluated for lack-of-fit (Section 2.7).

View ► Plot... ► Monte Carlo simulation (Ctrl+Shift+T).  This option produces a cell-distribution bar chart.

### 3.5.4.3  Data

The data sub-menu is shown in Figure 7 and has two choices:  Input data and model.  The model sub-menu is also shown in Figure 7 and has three choices: Relative resistance, Var(Y), and Monte Carlo simulation.
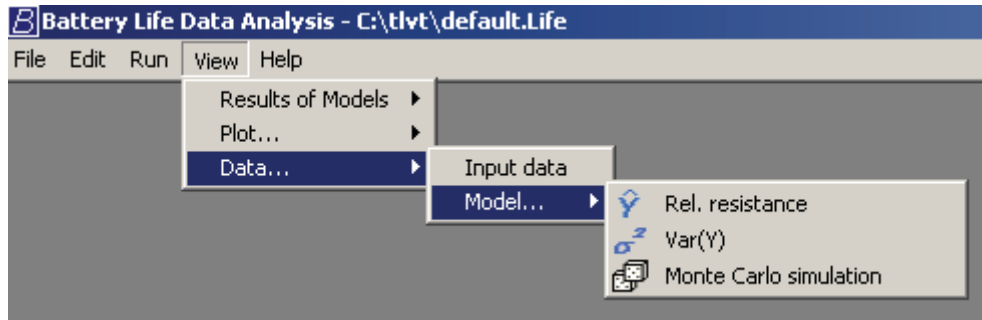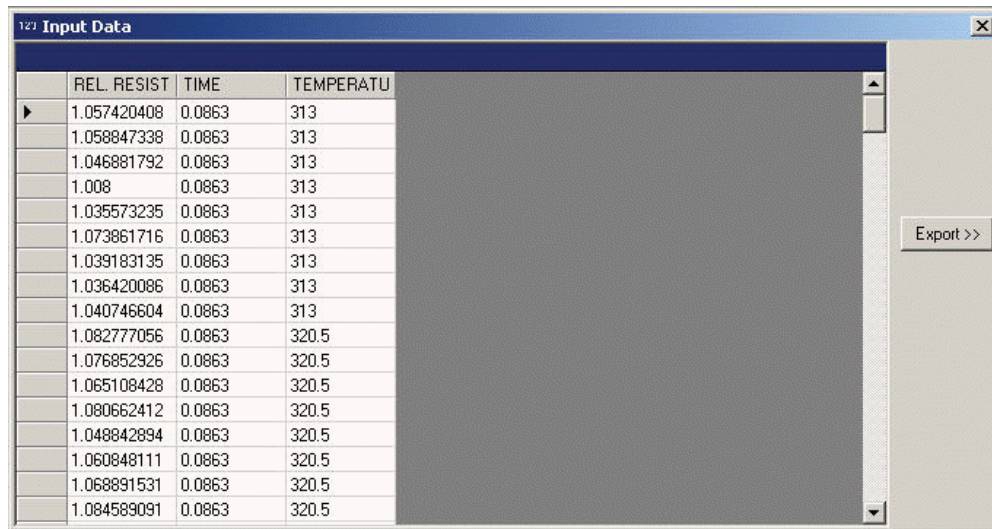


Figure 7.  View menu and data sub-menus

View ► Data ► Input Data.  This option displays the input data, and is useful for comparing the known input data to what the software read from the input file.  Figure 8 shows a sample data display with the proper format discussed in Section 3.4.  The data can also be sorted in ascending or descending order by simply clicking on the column title.  When the data are sorted by a particular column, a small arrow (▲or ▼) will appear that indicates the direction of the sort.  The data displayed in the grid can be exported to a comma-separated values (CSV) file by clicking on the "Export>>" button on the right hand side, as shown in Figure 8.

View ► Data ► Model... ► Rel. Resistance.  This option displays the degradation parameter data; in this case, it is the relative resistance.  The label will automatically adjust to the appropriate parameter being fit/modeled by the software.
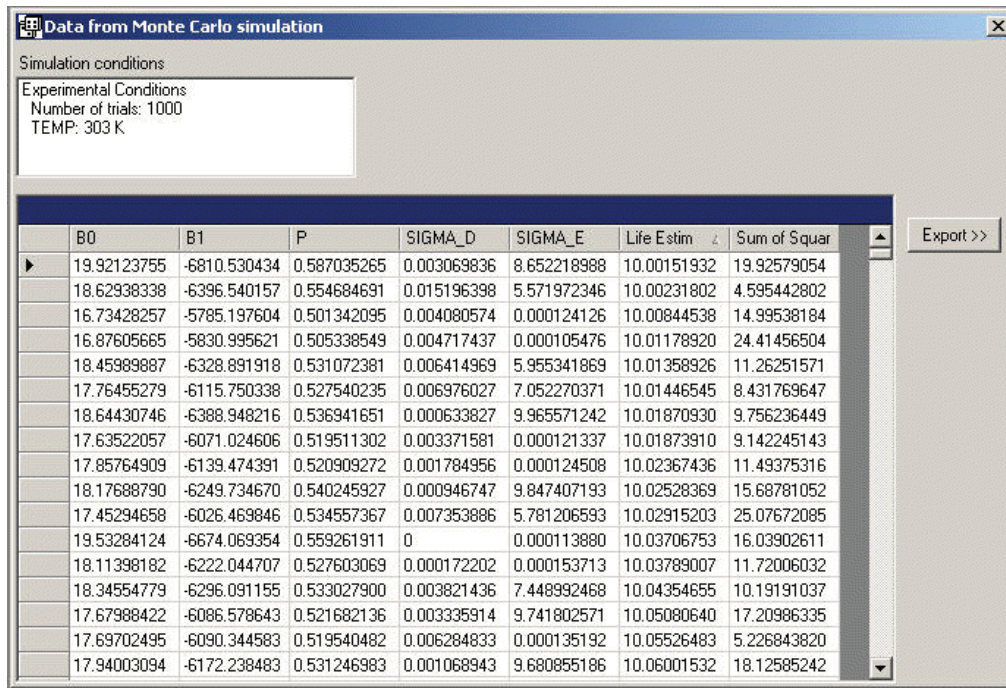
View ► Data ► Model... ► Var(Y).  This option displays the variance of the output data.

View ► Data ► Model... ► Monte Carlo simulation.  This option displays the results from the Monte Carlo simulation, including the parameter fits, error variances (*i.e.*, SIGMA_D is the variance of cell-to-cell effects and SIGMA_E is the variance of the measurement error), life projections, and the lack-of-fit sum of squares values generated for each simulation trial.  A sample Monte Carlo simulation output is shown in Figure 9. The data displayed in the grid can be exported to a comma-separated values (CSV) file by clicking on the "Export>>" button on the right hand side.



Figure 8.  Input data display



Figure 9.  Monte Carlo simulation data display

### 3.5.5  Help Menu

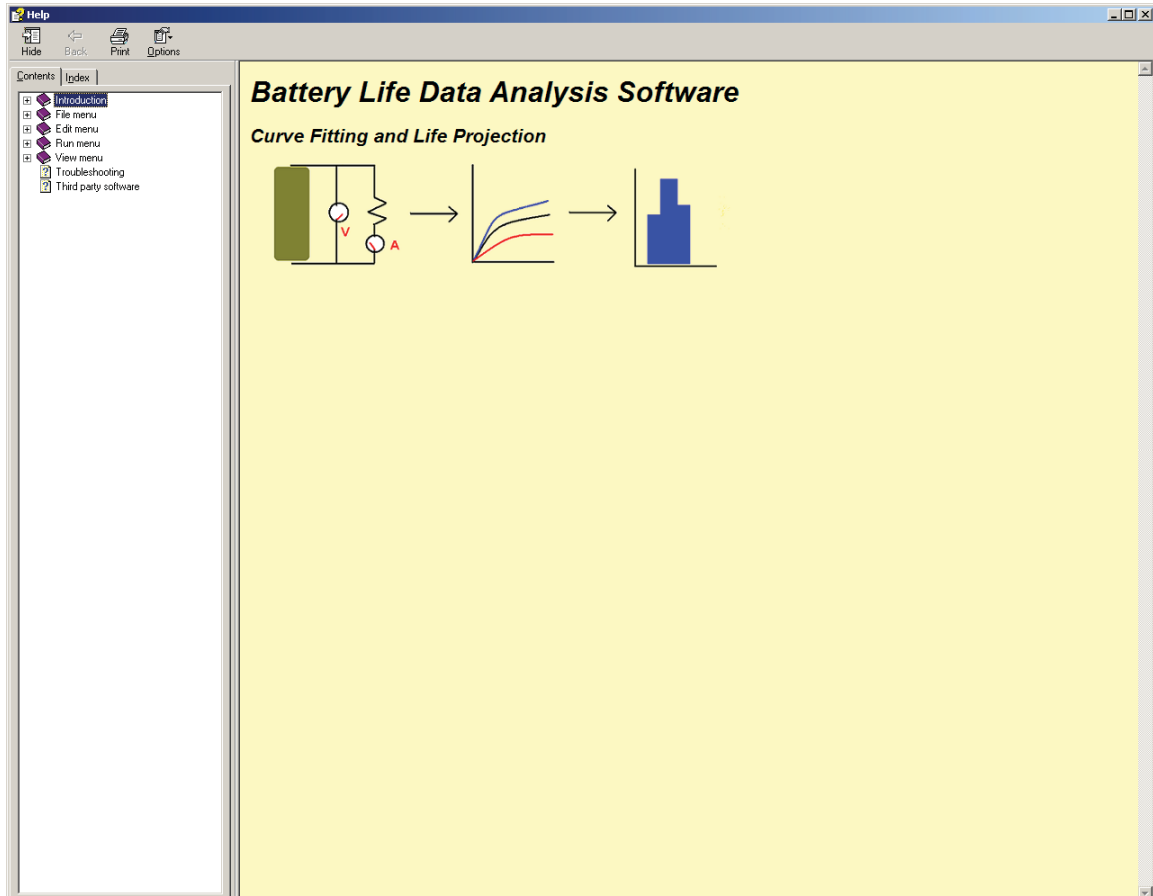The main help screen is shown in Figure 10 and provides a detailed description of the program's capability and navigation.



Figure 10.  Software help file

# 3.6  Using the Software

### 3.6.1  Creating or Opening an Existing Fit/Simulation File

Creating a new fit or simulation (File ► New ►) will immediately launch the appropriate wizard. When opening an existing file (File ► Open), the user will be prompted with a file-selection dialog box, similar to that shown below in Figure 11. The BLE software files are saved with a "*.Life" extension and contains information regarding what type of operation is to be done (fit, simulation or both) and the values of other parameters used in the program.
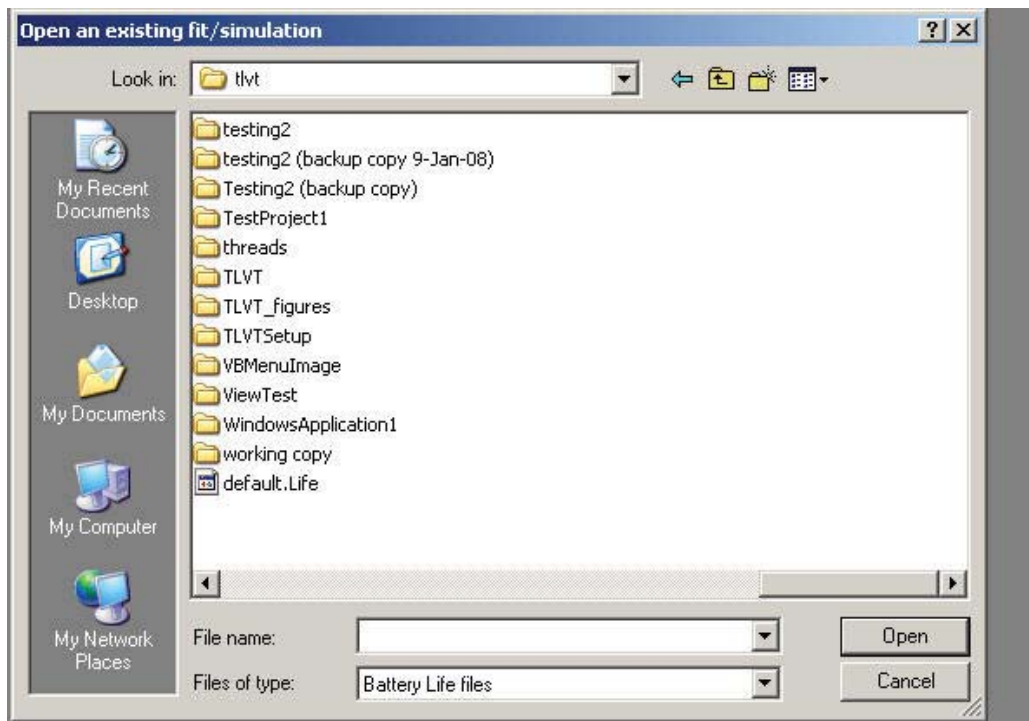


Figure 11.  File selection dialog box

### 3.6.2  Fitting and Simulation

Once a file has been opened or created, there are three options available in this software package for fitting and/or simulating the data:

1. Fit Only. This option solves for parameter values based on input data and the model equation.

2. Fit and Simulate. This option solves for parameter values and simulates life based on those fitted values.

3. Simulate Only. This option simulates life based on user-provided parameter values with the assumption that the fit was completed independent of this software package.

18

### 3.6.2.1 Fit Only

The program will prompt the user to identify the stress factor groups within the dataset to include in the fitting process. A sample dialog box is shown in Figure 12, with temperature (in Kelvin) identified as the stress factor of interest. The user may select/unselect the datasets indicated by the experimental stress condition(s) by clicking on the check boxes or on the Select All or Unselect All buttons. Click **OK** to proceed to the fitting process. The results of the fit can then be seen by clicking on the View menu choices (Section 3.5.4).
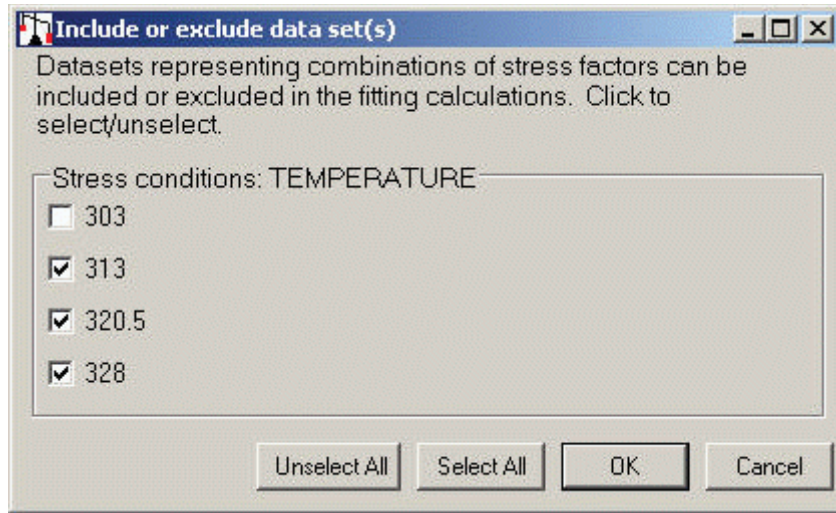


Figure 12. Include/exclude dataset(s) dialog box

### 3.6.2.2 Fit and Simulation

The program proceeds through the function described in Section 3.6.2.1, then prompts the user to run the simulation. The dialog then becomes similar to that found in the Fit Wizard discussion (Section 3.6.3). The results of the fit and simulation can then be seen by clicking on the View menu choices (Section 3.5.4).

### 3.6.2.3 Simulation only

The dialog is very similar to the description in Section 3.6.2.2, except that no fitting is performed and the parameters of the equations must be provided by the user. The results of the simulation can then be seen by clicking on the View menu choices (Section 3.5.4).

### 3.6.2.4 Life Confidence Limits

There are two types of life projections that can be performed with this program, one based on experimental data (Section 3.6.2.2) and the other based on parameters provided by the simulation only option (Section 3.6.2.3). The projection based on simulations only is primarily intended to establish a viable life verification experiment, so only the worst-case scenario (i.e., lower confidence limit) needs to be considered (Reference 4). The projection based on experimental data uses both an upper confidence

limit (UCL) and a lower confidence limit (LCL) to show the full range of uncertainty with regard to the expected life.  A typical display to change just the lower confidence limit is shown in Figure 13.  The life projection will then be calculated using the input value(s), then displayed on subsequent dialogue boxes and reports.
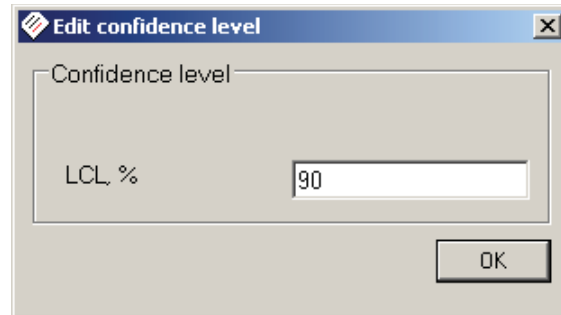


Figure 13.  Editing the lower confidence limit

### 3.6.3  Fit Wizard

The Fit Wizard assumes that time is an independent variable for all modeling equations.  The fitting process is as follows:

1. Specify the data to be fit

2. Associate variables to the data parameters

3. Specify the life equation and the corresponding transforms using the identified variables

4. Import test data from file

5. Identify stress factor combination(s) to include in the fit

6. Perform the fit

With each dialog box in the Fit Wizard, the user can move forward by clicking the "Next" button.  If all of the entries were not appropriately filled, the software will prompt the user for the missing information.  To update or modify a previous entry, press the "Back" button to return to the previous screen.

### 3.6.3.1  Defining the Data

The first step in the fitting process is to define the type of data that are to be used. Figure 14 shows the list of available dependent and independent variables in the software tool.  The dependent variables include resistance, capacity, energy, and power, and only one can be selected per fit.  The independent variables include time, temperature, state-of-charge (SOC), power, energy, capacity, discharge cutoff, and charge cutoff.  More than one independent variable can be chosen per fit, and they can be selected by clicking

on the variable name. For example, a total of three variables (one dependent, and two independent) have been identified for the fit in Figure 14.
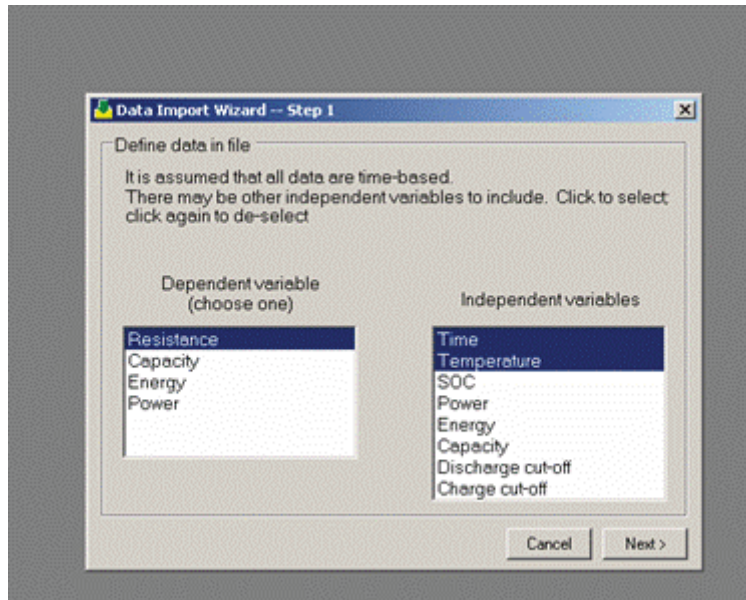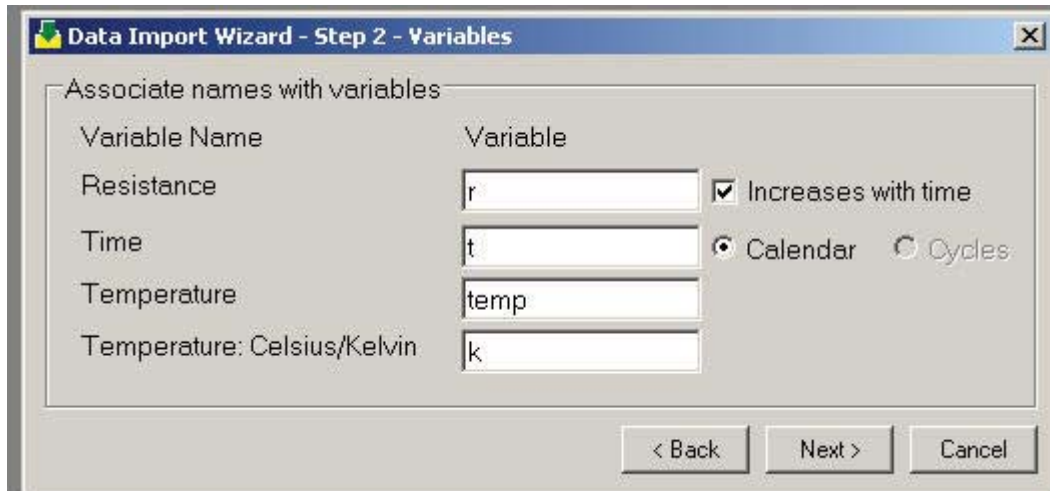


Figure 14. Typical display to define the data in a file

### 3.6.3.2 Associate Variables with the Definitions

Once the data have been defined, the next step is to assign variables to the parameters. Since there are three parameters identified in the example shown in Figure 14, the software will prompt the user to identify three algebraic variables, as shown in Figure 15. The algebraic variables can be more than one letter and the names are not case-sensitive (*e.g.*, "t" for time and "T" for temperature will not work). The first variable name in the list is the dependent variable (*e.g.*, resistance). The software will not associate any physical process with the name and, hence, does not implicitly know if the dependent variable increases or decreases with time. Therefore, the "Increases with time" box to the right of the dependent variable text box should be checked when appropriate (*e.g.*, it is checked in the case of resistance, but wouldn't be in the case of capacity or power loss).

All temperatures in the fitting and simulation processes must be in Kelvin. If the data input already has temperature in Kelvin, put a "k" in the "Temperature: Celsius/ Kelvin" text box. Otherwise, place a "c" in the text box, and the software will automatically perform the conversion to Kelvin.

At this stage in the software development, only calendar-life data in units of time (*e.g.*, weeks, months, years, etc.) are considered for fitting. At some future date, cycle life data will also be implemented in the software tool as well. Until then, the "Cycles" option is unavailable.

Figure 15. Typical display to associate data and variables

### 3.6.3.3  Specify the Equations

The life and transform equations must be entered next. The non-linearized form of the life equation is entered in the upper text box marked "Modeling equation," and must contain all of the independent variables identified in the previous step (Section 3.6.3.2). The form of the equations must be of the appropriate mathematical operations and functions.  The "Help" button will identify the available algebraic symbols, as shown in Figure 16.  Do not include an equal sign (*i.e.*, "=") at the beginning of the expression.



Figure 16. Available algebraic functions

The life equation must be easily linearized through simple mathematical operations (Section 3.3). The terms in the linear form of the equation are to be entered in the group of text boxes labeled as "Transform equations." The number of transform equations will depend on the number of independent variables identified in Section 3.6.3.1.

Figure 17 shows the dialog box for specifying the equation and transforms using the example variables that have been identified (*e.g.*, resistance, time, and temperature). The equation is based on the default model given in Section 2.2. The "Default Model" button in the dialogue box will automatically fill in the text boxes with the appropriate model and transforms using the variable labels identified in Section 3.6.3.2. In the example, the life equation is $R=1+\exp(b_0+b_1/temp)t^z$. The linearized form is, therefore, $\ln(R-1)=b_0 + b_1/temp + z\ln t$. The terms containing the three variables entered above are entered in the boxes; the fitting constants, $b_0$, $b_1$ and $z$, are not entered. The order in which the transform equations are entered does not matter. The equations are not case sensitive and there is no limit to the number of terms in them (Section 3.3).
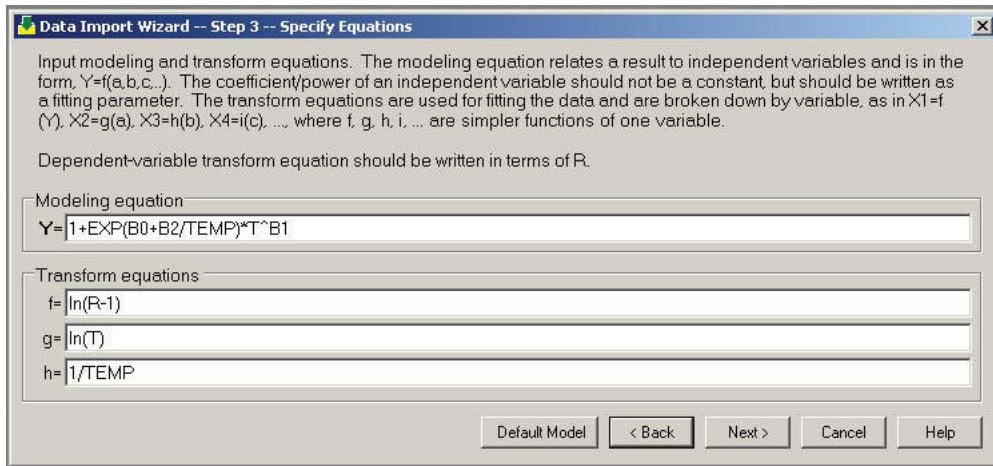


Figure 17. Typical display to specify equations

### 3.6.3.4 Import Test Data

The next step in the Fit Wizard is to import data from an Excel workbook or comma-separated value file. The data must be arranged in columns and normalized to time $t = 0$ (see Section 3.4 for data formatting requirements). Additionally, if the relative change between measurements is greater than 50% (increase or decrease), the error model strategies will not be applicable and the user will then see a warning to this effect and have the option of continuing or not. Figure 18 shows a typical display for file selection. Only one file can be selected at a time.
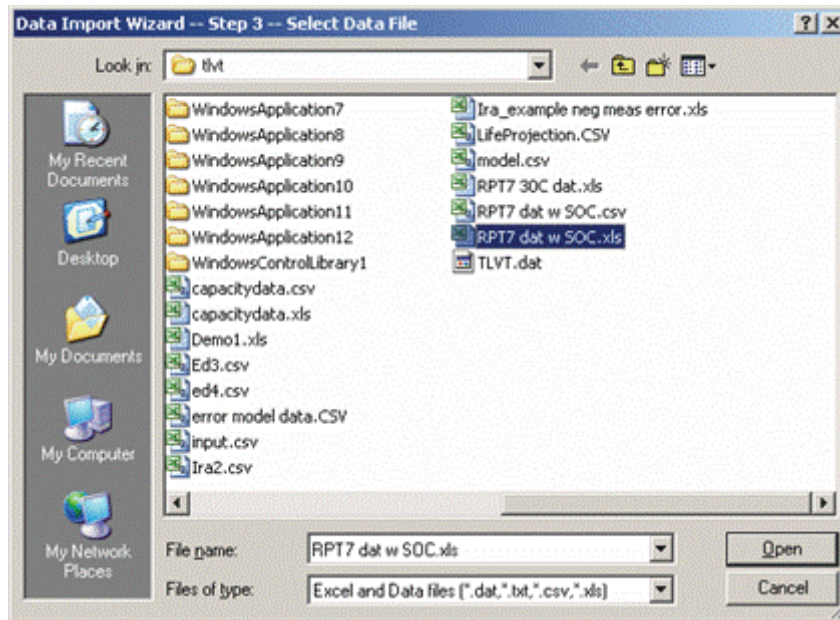
Figure 18. Typical display to select a data file

Once a file has been selected, the wizard will produce a display similar to that shown in Figure 19. In this example, an Excel workbook was selected as the data file, so the worksheets names are shown on the left side of the display. Clicking on the worksheet with the data of interest produces a display similar to Figure 20. If an error was made in the file selection, click **Open another file...** to change source data. A similar display will be shown if data are read from a comma-separated-values file.



Figure 19. Typical display to open an Excel file



Figure 20. Typical display to select columns from a particular worksheet

24

In Figure 20, the "UseThis" worksheet is selected. In this case, the program displays the contents of the first two rows of the selected sheet in the preview window, allowing for easier column selection. The variables are associated with the corresponding data 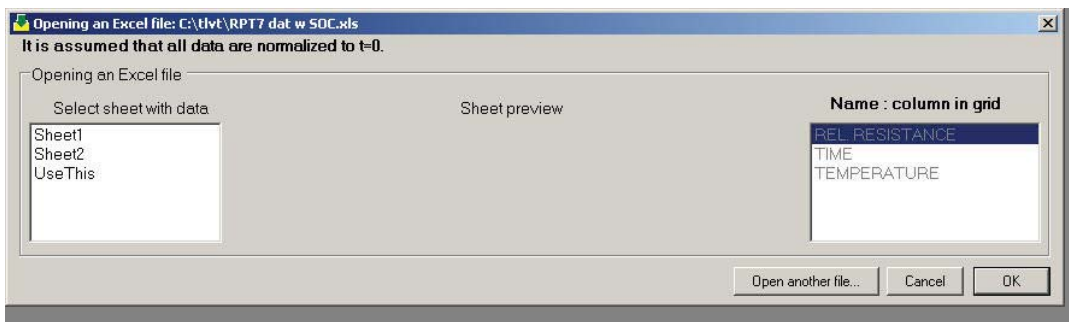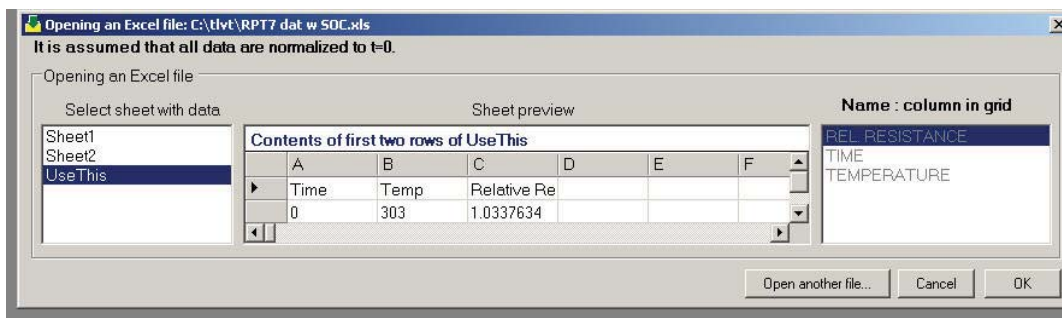by clicking on the first cell of the appropriate columns. For the example shown in Figure 20, the mouse-click sequence would be columns C-A-B, corresponding to the list of variables identified on the right-hand side. Once a variable has been assigned, the column location is identified in the right window (*e.g.* "TIME" will change to "TIME: A"). If a mistake is made, either continue choosing columns until display cycles back to where the error is or click on the incorrect name-column combination on the right and then click on the correct column. It is assumed that the time data are in calendar years and that the temperatures are in Celsius or Kelvin. If the temperatures are in Celsius, the program will convert them to Kelvin (Section 3.6.3.2).

### 3.6.3.5 Select Stress Factor Combinations

After the data are imported, they will be analyzed and grouped according to the stress factors previously identified (see Figure 14). There is also an opportunity to include or exclude a set of stress conditions if the data are not well-behaved. Figure 21 shows the display for this step using temperature as an example stress condition. Following this step, the software proceeds to the fit, and the results can be displayed as shown in Section 3.7.
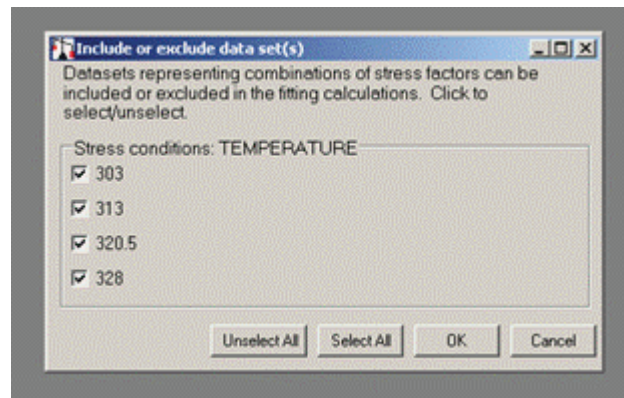


Figure 21. Selecting which stress factors to include in the fit calculation

### 3.6.4 Simulation Wizard

Simulation is performed to calculate a life projection based on an assumed model, and, if experimental data are available, to judge the lack of fit between the test data and the model. As with the Fit Wizard (Section 3.6.3), the Simulation Wizard assumes time is an independent variable for all modeling equations. The simulation process is shown below.

1. Define the data to be used

2. Associate variables to the data parameters

3. Specify the life equation and the corresponding transforms

4. Enter simulation parameters

5. Specify simulation conditions

6. Perform the simulation

With each dialog box in the Simulation Wizard, the user can move forward to the next step by clicking the "Next" button. If not all of the entries were appropriately filled, the software will prompt the user for the missing information. To update or modify a previous entry, press the "Back" button to return to the previous screen. The first three steps of the Simulation Wizard are the same as the Fit Wizard, and are discussed in Sections 3.6.3.1 through 3.6.3.3; Steps four through six are described below.

### 3.6.4.1 Enter Simulation Parameters

The parameters affecting the simulation results are entered using the form shown in Figure 22. Most of the fields will be automatically filled in by the software program based on the results from the Fit Wizard (Section 3.6.3). The user must enter a value for the percent-change of the degradation parameter at end-of-life (EOL). A typical EOL value is 30% degradation. The other fields can also be adjusted as necessary, though it is not recommended if the user wants the simulation to replicate the actual experiment as closely as possible. For example, if the measurement error is independently determined (Section 2.3.2), then this value can be entered in the appropriate field. The initial parameter values (*e.g.*, B2, B1, and B0 in this case) should be guesstimated based on expected performance. For example, if the model generally assumes a square root of time dependence, the value for "B1" in Figure 22 should be close to 0.5.

Figure 22. Display allowing entry of simulation parameters

### 3.6.4.2 *Specify Simulation Conditions*

The experimental conditions (temperature, duration, number of cells, etc.) are then specified for all groups, as indicated in Figure 23. There is no limit to the number of groups that can be used, but the total number of cells in the simulation is limited to 500. The Monte Carlo simulation parameters for life prediction are also entered at this stage. These values include the temperature (in Kelvin), the upper and lower confidence limits (only the LCL is needed if simulating results without experimental data), and the number of simulation trials. Click on **OK** to perform the simulation and life prediction.



Figure 23. Experimental conditions and Monte Carlo parameter entry

## 3.7  Reports and Displays

### 3.7.1  Model Fit Results

The results from the model fit can be found by clicking on View ► Results of Models ► Fitting (Ctrl+Shift+G). An example result is shown in Figure 24, and includes the model equation, the names and variables used, the values of the fitting parameters, the estimate of errors, and lack of fit information. The numbers in parentheses given for each of the fitting parameters are the bootstrap standard errors determined from the Monte Carlo simulations. The values given in the error model section, $\sigma_\varepsilon^2$ and $\sigma_\delta^2$, represent the variances of the measurement error and cell to cell effects, respectively. They are derived from the regression of the variance of the experimental dependent variable, Y, on $(\hat{\mu} - 1)^2$, where $\hat{\mu}$ is the predicted value of Y. The variance of the dependent variable, Var(Y), and $\hat{\mu}$ are computed for each experimental condition (*e.g.*, temperature/time) to form the data set used in this regression. Alternatively, the measurement error will reflect the independently determined value, as discussed in Section 2.3.2.

The value of the lack-of-fit statistic ($SS_{LOF}$) obtained with the actual data is compared with the empirical cumulative distribution function (CDF) of the $SS_{LOF}$ values obtained via the simulation trials to assess model inaccuracy. If the lack-of-fit statistic ($SS_{LOF}$) obtained with the actual data exceeds the $(1 - \alpha) \cdot 100^{th}$ percentile of the

simulated $SS_{LOF}$ values, then it can be concluded that the model is inaccurate. Normally, a small value for $\alpha$ is selected (*e.g.*, 0.05). In practice, the CDF is evaluated at the value of the lack-of-fit statistic ($SS_{LOF}$) obtained with the actual data. If the evaluated CDF point exceeds $(1-\alpha)$, then it is concluded that there is evidence for lack of fit. Pressing the "View CDF" button will display the empirical CDF curve and the evaluated CDF point as shown in Figure 25. The empirical CDF curve may take some time to display fully.
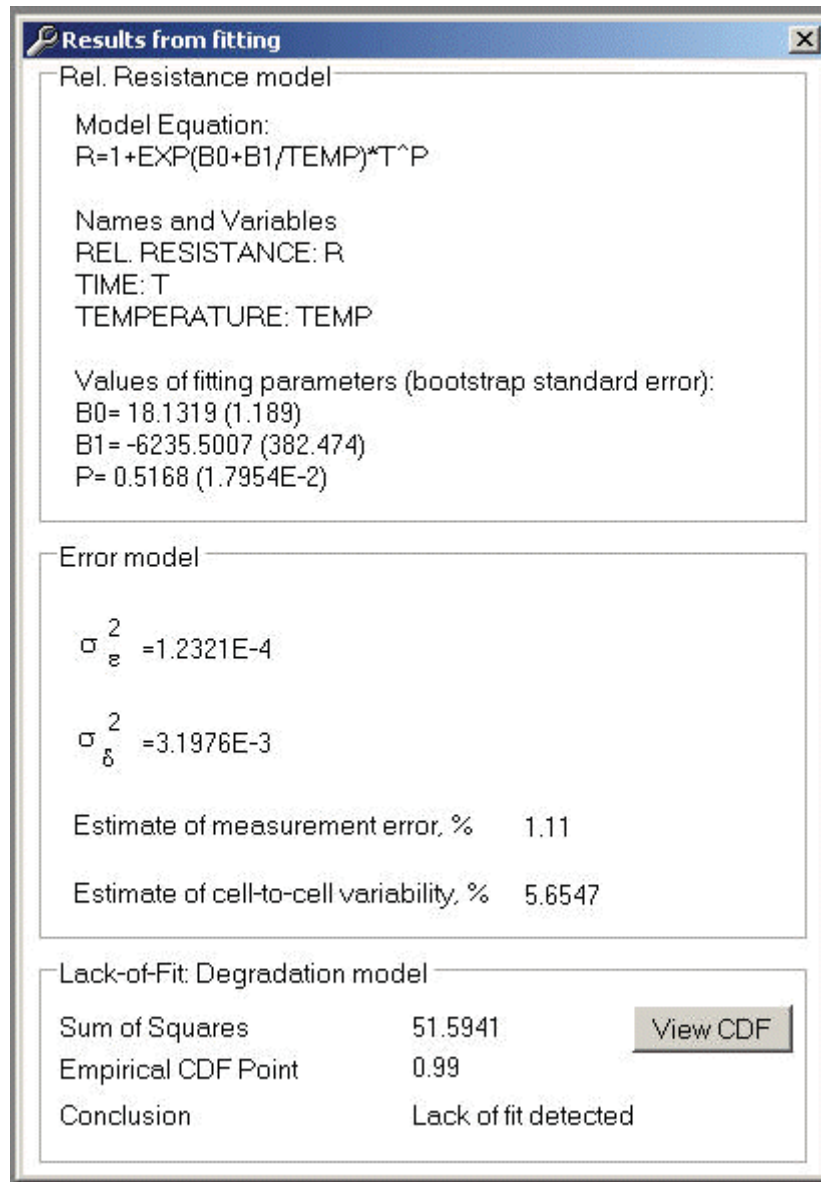


Figure 24. Example results from fitting display
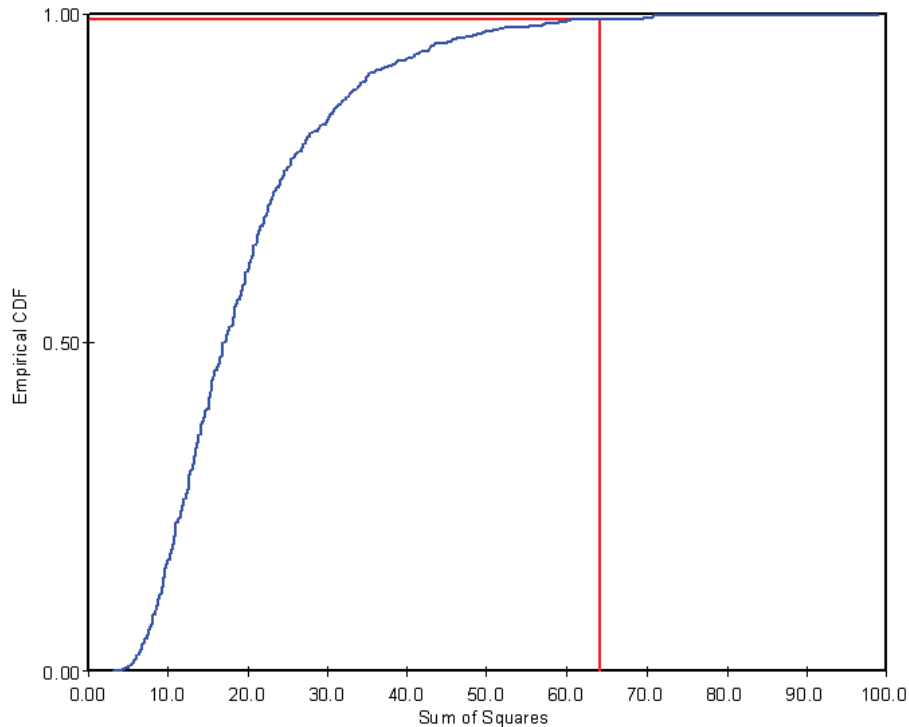
Figure 25.  Example CDF curve

## 3.7.2  Monte Carlo Simulation Results

### 3.7.2.1  Life Estimation

The results of the Monte Carlo simulation can be found by clicking on View ►
Results of Models ► Monte Carlo simulation (Ctrl+Shift+S).  Figure 26 shows an
example of life estimation when using experimental data (*i.e.*, a life estimate with an
upper and lower confidence limit).  If a life estimate is based only on simulated data
derived from the input model parameters, the resulting prediction will typically be
displayed as shown in Figure 27, with only the LCL life projection.



Figure 26.  Life estimate display using experimental data

29

Figure 27.  Typical life projection result display from Monte Carlo simulation

### 3.7.2.2  Data Plots

As mentioned in Section 3.5.4.2, there are five data plots available in the View ▶ Plot... ▶ menu option.  A sample fitting results plot is shown in Figure 28 a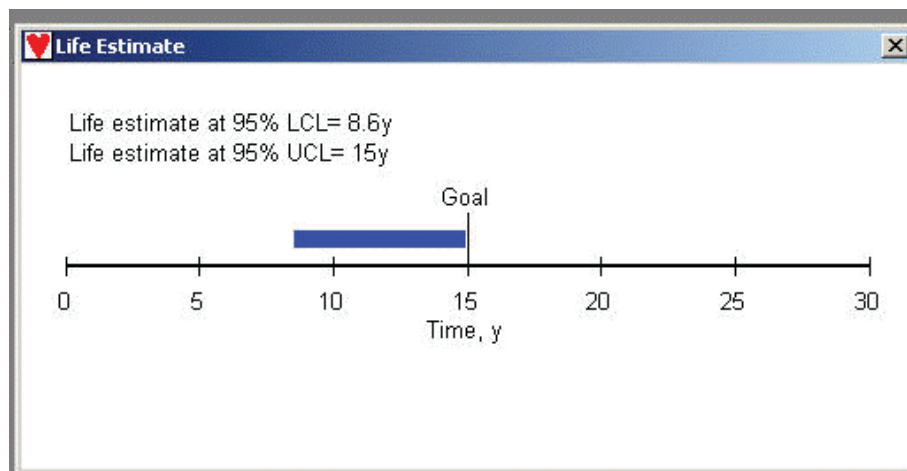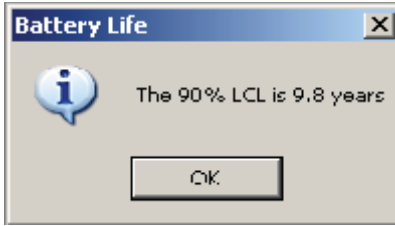nd shows the experimental data and the resulting model fit.  The numbers on the right of the model output indicate the test temperature (in Kelvin).  An example plot of the error model results (*i.e.*, Var(Y) vs. ($\hat{\mu}$ -1)²) is shown in Figure 29, where $\hat{\mu}$ is the estimated performance.  A plot of ($\bar{Y}$ - $\hat{\mu}$) vs. $\hat{\mu}$ is shown in Figure 30, where $\bar{Y}$ is the average measured performance.  A plot of $\hat{\mu}$ vs. $\bar{Y}$ is shown in Figure 31, including the 45° line that indicates a perfect fit.  The simulated data in Figure 31 shows evidence of lack-of-fit, as concluded in the summary display (Figure 24).  Finally, a sample cell-distribution bar chart determined from the Monte Carlo simulation is shown in Figure 32.  This particular plot shows that over 50% of the cells have a predicted life of about 12 years or more, and that the highest population group has a predicted life of 11 years.
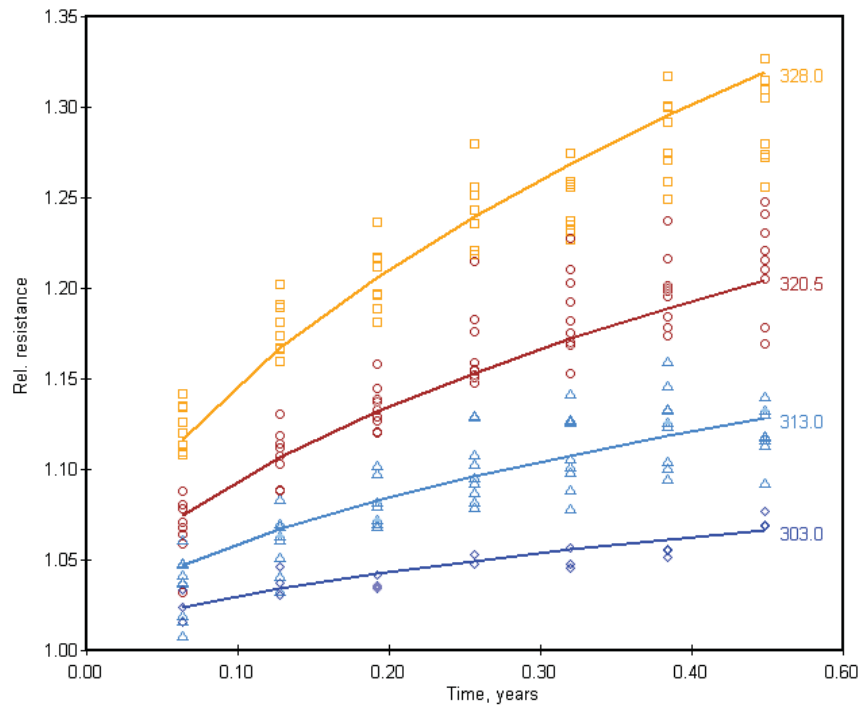


Figure 28.  Experimental data and fitted curves
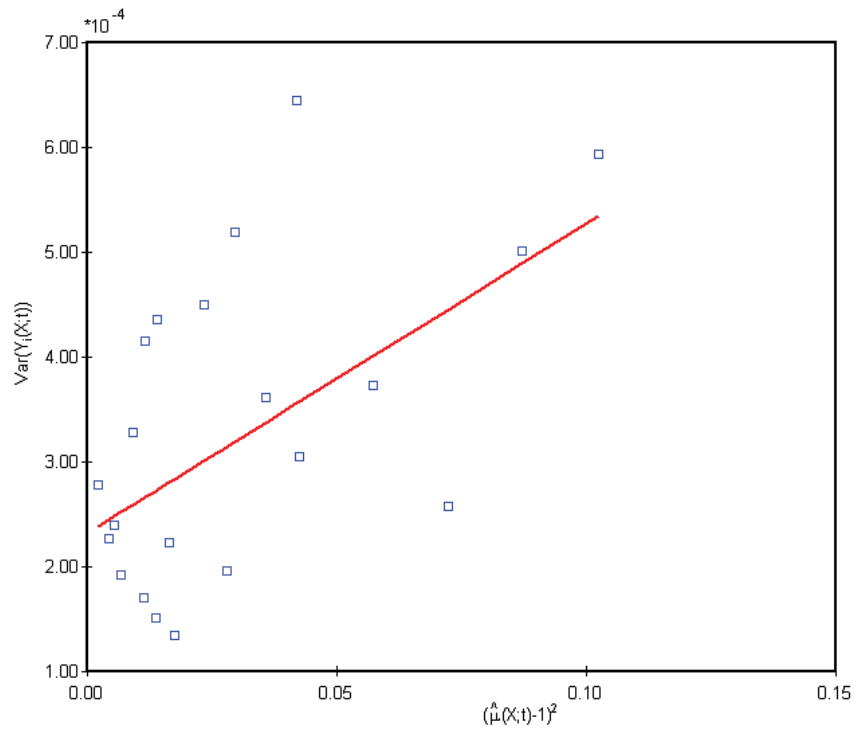
Figure 29. Error model results
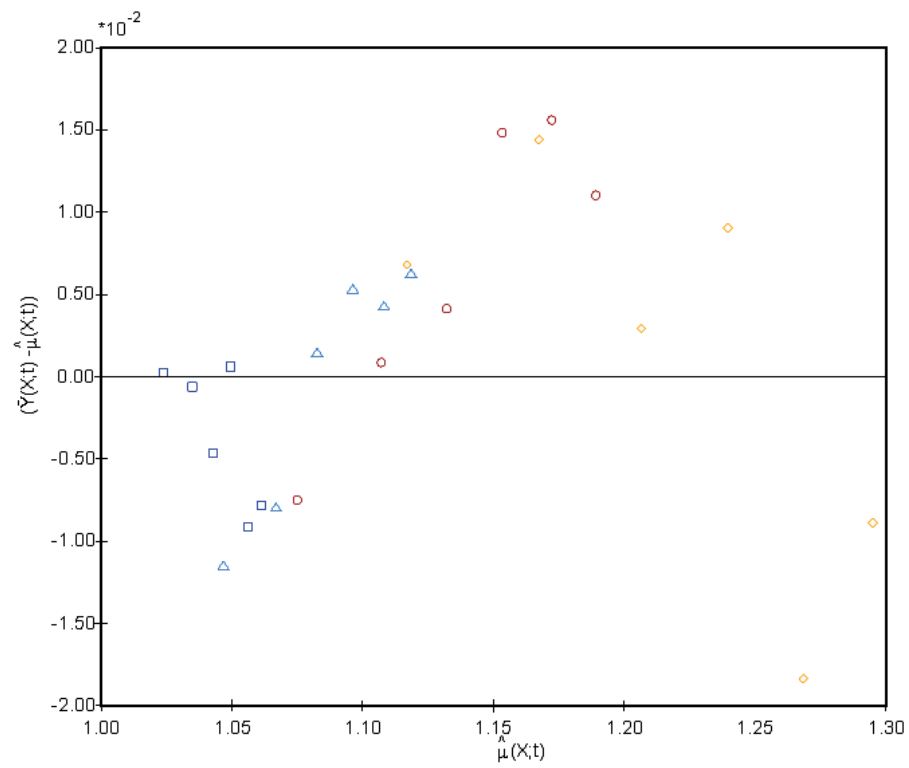


Figure 30. Difference between average and predicted performance vs. predicted performance
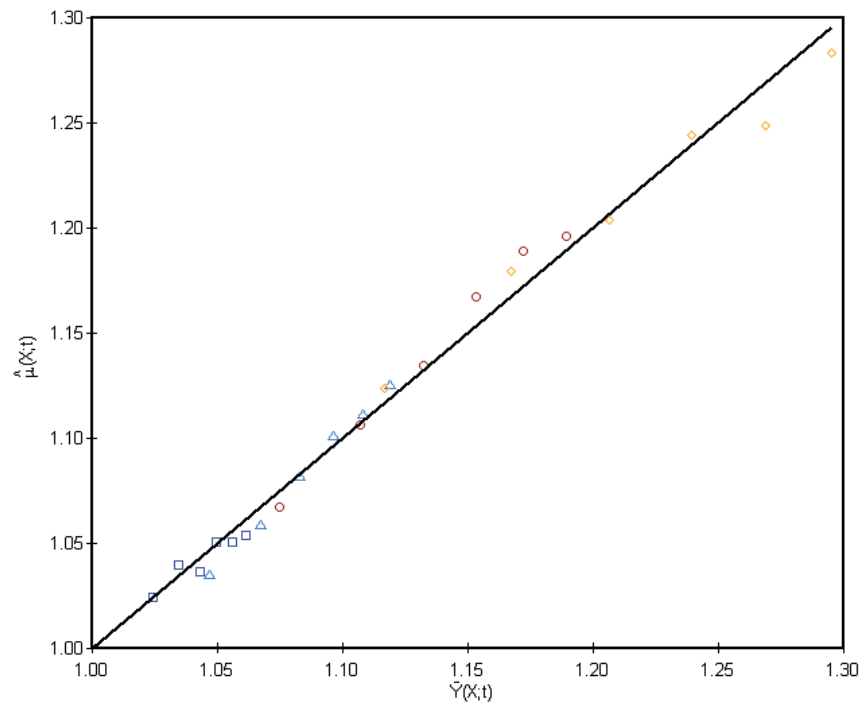
Figure 31.  Predicted performance behavior vs. the average measured performance
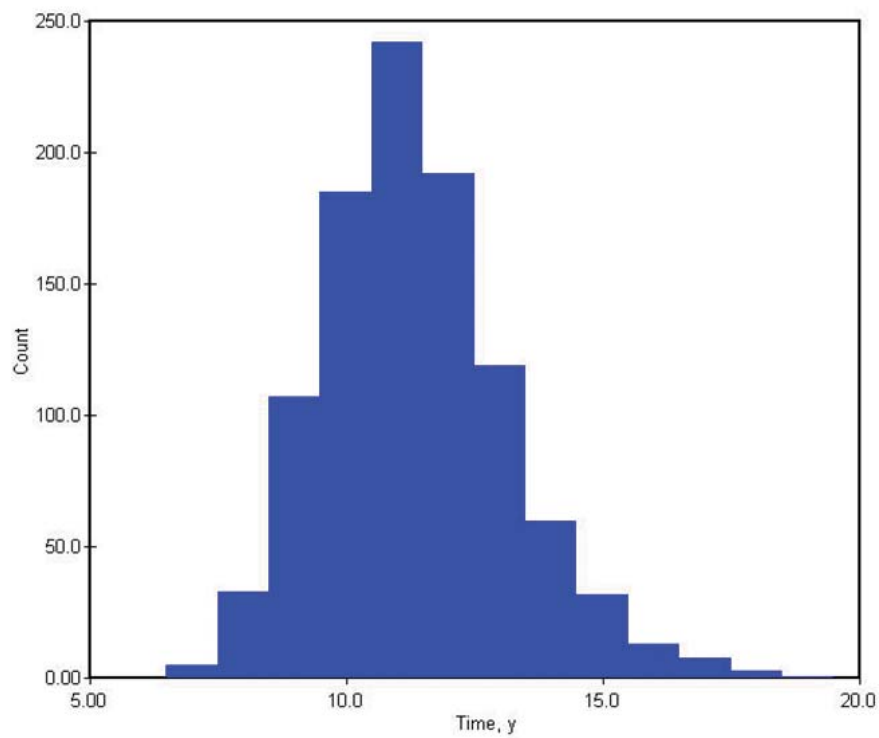


Figure 32.  Cell distribution bar chart

### 3.7.3  Report Preview

A preview of the results assembled by the software can be found by clicking File ► Preview Report (Section 3.5.1).  A sample report is shown in Figure 33.  This is useful in viewing the report prior to printing it on paper.  The '<Prev' and 'Next>' button allow navigating between and among a multi-page report.
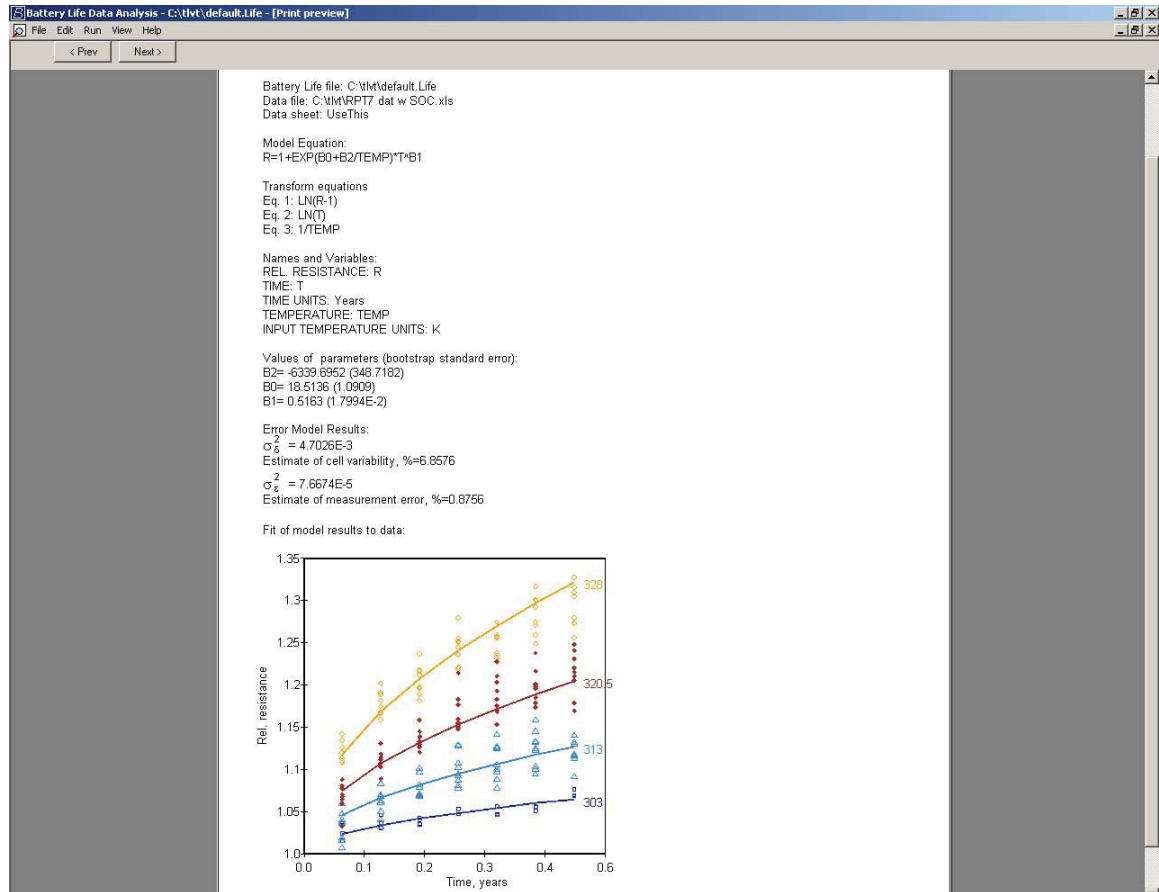


Figure 33.  Report print preview

# 3.8   Troubleshooting

The following section identifies known issues with the software development as of the publication of this manual, and their associated solutions.

## 3.8.1   Text-entry boxes and/or buttons are not in the correct positions or appear to be crowded

**Solution:**  Change the settings for the display.  Close the application.  Right click on the desktop and select properties.  Click on the tab that says Settings.  Click on the button that says Advanced.  Change the DPI setting in the Display group box to 96 DPI. Click OK twice.

## 3.8.2   The application is running on a network and an error occurs when trying to open a Life file

**Solution:**  A message box similar to the one shown in Figure 34 should appear. This error occurs only when the application is running on an intranet.  Download the .NET 2.0 SDK from Microsoft and follow the install instructions given in http://technet.microsoft.com/en-us/library/bb742442.aspx.  Reboot the computer.  Click Start -> Programs -> Administrative Tools -> Microsoft .NET Framework 2.0 Configuration.  Click on the 2.0 Configuration snap-in followed by OK.  Click on Configure Code Access Security Policy -> Adjust Zone Security -> Make changes to this computer -> OK -> Local intranet.  Move slider to Full Trust.  Click Next -> Finish. Close the configuration tool window.  The application should run normally now.
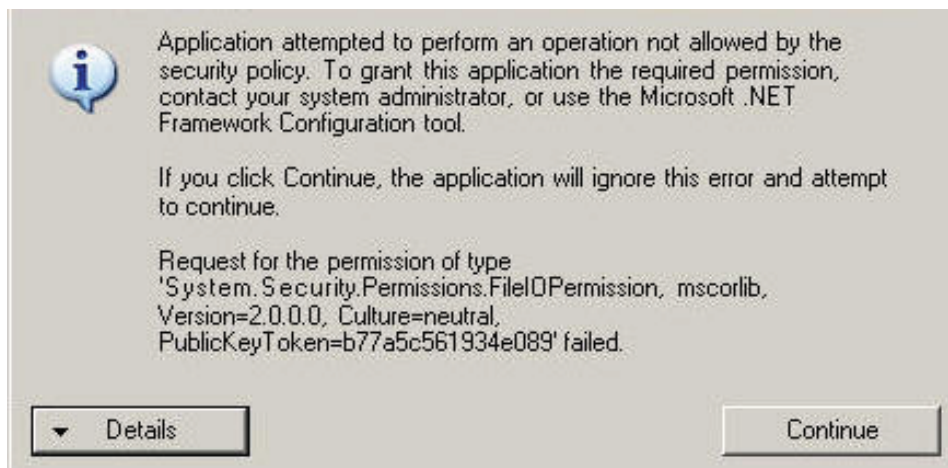


Figure 34.  Network error box

# 4 REFERENCES

1. *Battery Test Manual for Plug-In Hybrid Electric Vehicles*, INL/EXT-07-12536, 2008.

2. *FreedomCAR Battery Test Manual for Power-Assist Hybrid Electric Vehicles*, DOE/ID-11069, 2003.

3. *FreedomCAR 42-volt Battery Test Manual*, DOE/ID-11070, 2003.

4. *FreedomCAR Ultracapacitor Test Manual*, DOE/NE-ID-11173, September 2004.

5. *Advanced Technology Development Program for Lithium-Ion Batteries: Battery Technology Life Verification Test Manual*, INEEL/EXT-04-01986, 2005.

6. E. V. Thomas, I. Bloom, J. P. Christophersen, and V. S. Battaglia, "Statistical methodology for predicting the life of lithium-ion cells via accelerated degradation testing," *J. Power Sources,* **184**, 312-317 (2008).

7. *Uncertainty Study of INEEL EST Laboratory Battery Testing Systems*, INEEL/EXT-01-00505, Vol. 1, December 2001.

8. *Uncertainty Study of INEEL EST Laboratory Battery Testing Systems*, INEEL/EXT-01-00505, Vol. 2, March 2003.

9. Holland, P. W. and Welsch, R. E., "Robust Regression Using Iteratively Reweighted Least-Squares," *Communications in Statistics: Theory and Methods*, A6, 1977, pp. 813-827.

10. Iglewicz, B., "Robust Scale Estimators and Confidence Intervals for Location," *Understanding Robust and Exploratory Data Analysis,* D. C. Hoaglin, F. Mosteller, and J. W. Tukey, eds., New York: Wiley, p. 417, 1983.

11. B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, London, U.K., 1993.

# Appendix A

# Methods for Estimating Model Parameters

# Appendix A

# Methods for Estimating Model Parameters

## 1   INTRODUCTION

This appendix summarizes the statistical methodology used for estimating the parameters associated with the degradation and error models.  In both cases a linear model that can be represented by the form shown in Equation (A1) is assumed, where $Z$ is a response variable, the $X_i$'s represent $M$ explanatory variables, and the $\beta_i$'s are model parameters which are to be estimated.  To achieve this linear form, it may be necessary to transform the natural response and/or one or more of the natural explanatory variables to a linear form.  For example, the response could be transformed as shown in Equation (A2), where $Y$ is the underlying performance metric (*e.g.*, relative resistance).  Other transformations could also be applied to explanatory variables, such as the inverse temperature.

$$Z = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \ldots + \beta_M \cdot X_M \tag{A1}$$

$$Z = \log(Y - 1) \tag{A2}$$

For estimating model parameters ($\beta_i$), it is assumed that there are $N$ observations, each containing the observed value of the response variable, $Z$, and the associated values of the explanatory variables, $X_i$ (*e.g.*, inverse temperature).  Thus, the data consist of a set of observations, as shown in Equation (A3).  The model parameters are then estimated by using a robust linear regression procedure (Reference 9).  The purpose of using a robust regression procedure rather than ordinary least squares is to reduce the influence of anomalous data on the parameter estimates.

$$\left\{ \left( Z(i); X_1(i), X_2(i), \ldots, X_M(i) \right) : i = 1, 2, \ldots, N \right\} \tag{A3}$$

An illustrative example of this methodology is provided in Reference 6.

## 2   ROBUST LINEAR REGRESSION

Two specific cases for estimation are considered.  In the first case, the model contains an intercept term ($\beta_0$).  Here, the initial step is to compute the average and standard deviation of each of the explanatory variables, $X_i$.  This yields $\left\{ \overline{X}_1, \overline{X}_2, \ldots, \overline{X}_M \right\}$ and $\left\{ S_1, S_2, \ldots, S_M \right\}$, respectively.  These explanatory variables are then mean-centered and scaled as shown in Equation (A4).  Therefore, in the case with an intercept, the design matrix ($D$) has a size of "$N$ by $M+1$", and is of the form shown in Equation (A5).

$$X_j^*(i) = \frac{X_j(i) - \overline{X}_j}{S_j} \tag{A4}$$

$$D = \begin{bmatrix} 1 & X_1^*(1) & \cdots & X_M^*(1) \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ 1 & X_1^*(N) & \cdots & X_M^*(N) \end{bmatrix} \qquad (A5)$$

In the second case for estimation, the model does not contain an intercept. Here, the explanatory variables are not mean-centered and/or scaled, and the design matrix ($D$) has a size of "$N$ by $M$", and is of the form shown in Equation (A6).

$$D = \begin{bmatrix} X_1(1) & \cdots & X_M(1) \\ \vdots & \cdots & \vdots \\ X_1(N) & \cdots & X_M(N) \end{bmatrix} \qquad (A6)$$

Once the design matrix has been determined, initialize a weighting matrix ($W$) of size "$N$ by $N$" with the identity matrix, and set $Z = (Z(1), Z(2), \ldots, Z(N))^T$, where $Z$ is the response variable of the form shown in Equation (A2). Next, repeat the following matrix computations three times:

$$b = \left( D^T \cdot W \cdot D \right)^{-1} \cdot D^T \cdot W \cdot Y \qquad (A7)$$

$$R = Z - D \cdot b \qquad (A8)$$

$W$ = diagonal matrix with elements from $BIWEIGHT(R_1, R_2, \ldots, R_N)$ $\qquad$ (A9)

The values for $b$ in Equation (A8) are determined from Equation (A7) and the corresponding explanatory variables. In the first case (with an intercept), the estimated model parameters are given by Equations (A10) and (A12). In the second case (without an intercept), the estimated model parameters are given by Equation (A12).

$$\hat{\beta}_0 = \left( b_0 - b_1 \cdot \frac{\overline{X}_1}{S_1} - b_2 \cdot \frac{\overline{X}_2}{S_2} - \ldots - b_M \cdot \frac{\overline{X}_M}{S_M} \right) \qquad (A10)$$

$$\hat{\beta}_i = \frac{b_i}{S_i} \text{ for } i = 1: M \qquad (A11)$$

$$\hat{\beta}_i = b_i \text{ for } i = 1: M \qquad (A12)$$

The *BIWEIGHT* function in Equation (A9) produces an *N-dimensional* output of weights from an *N-dimensional* input. It is based on Tukey's biweight function with c = 6 (Reference 10). Letting $BIWEIGHT\{R_1, R_2, \ldots, R_n\}$ represent the biweight function and its input arguments, determine the weighting function as follows:

1. Compute the median absolute value (*MAV*) of the $R_i$'s determined from Equation (A8)

2. Compute standardized values of $R_i$'s as shown by Equation (A13)

3. $WEIGHT_i = 0$ if $Z_i \geq 1$ for $i = 1:n$

4. $WEIGHT_i = \left(1 - U_i^2\right)^2$ if $Z_i < 1$ for $i = 1:n$

5. Return weights (*W*) to Equation (A9). The updated diagonal elements of *W* are given by the values $\{WEIGHT_1, WEIGHT_2, \ldots, WEIGHT_n\}$.

$$U_i = \left. Z_i \middle/ \left(c \cdot MAV\right) \right. \tag{A13}$$

## 2.1 Illustration of Linear Model Forms

The default degradation model has the linearized form shown in Equation (A14). This can be put into the form of Equation (A1) using Equation (A2) and the corresponding transformations shown in Equations (A15) through (A17), where the $X_i$'s are generic explanatory variables as defined by Equation (A3).

$$\log(\mu(T;t) - 1) = \beta_0 + \beta_1 \cdot \frac{1}{T} + \rho \cdot \log(t) \tag{A14}$$

$$X_1 = \frac{1}{T} \tag{A15}$$

$$X_2 = \log(t) \tag{A16}$$

$$\beta_2 = \rho \tag{A17}$$

The default error model has the linearized form given in Equation (A18), with the corresponding transformations shown in Equations (A19) through (A22).

$$Var\left(Y_i(T;t)\right) \approx \sigma_\delta^2 \cdot \left(\mu(T;t) - 1\right)^2 + \sigma_\pi^2 \tag{A18}$$

$$Z = Var\left(Y_i(T;t)\right) \tag{A19}$$

$$X_1 = \left(\mu(X;t) - 1\right)^2 \tag{A20}$$

$$\beta_0 = 2 \cdot \alpha^2 = \sigma_\pi^2 \tag{A21}$$

$$\beta_1 = \sigma_\delta^2 \tag{A22}$$

# 3 POSSIBLE COMPLICATIONS WHEN ESTIMATING ERROR MODEL PARAMETERS

Due to statistical fluctuations in the data, use of the estimation procedure might result in $\hat{\alpha}^2 < 0$ (i.e., $\hat{\beta}_0 < 0$) or $\hat{\sigma}_\delta^2 < 0$ (i.e., $\hat{\beta}_1 < 0$). Such negative estimates of variances can and should be regarded as nonsensical and the parameter estimates should be modified as discussed below.

## 3.1 Case 1: $\hat{\alpha}^2 < 0$

In this case, the hypothesized measurement error was not detected. First, it is recommended that an alternative non-negative estimate for $\alpha^2$ be specified (*i.e.*, $\hat{\alpha}^2(alt)$). This estimate could be based on an independent assessment of the measurement error (Section 2.3.2, References 7-8). If the measurement error variance is believed to be vanishingly small when compared to $\sigma_\delta^2$, then it might be prudent to consider $\hat{\alpha}^2(alt) = 0$.

Once the alternative estimate of $\hat{\alpha}^2$ is available, $\hat{\sigma}_\delta^2$ must be re-estimated. This can be accomplished by first re-parameterizing the error model as shown in Equation (A23).

$$Z = Var\left(Y_i\left(X;t\right)\right) - 2 \cdot \hat{\alpha}^2\left(alt\right) = \beta_1 \cdot \left(\mu\left(X;t\right) - 1\right)^2 \tag{A23}$$

For purposes of regression, each observation again consists of $Z$ and $X_1$, where $X_1$ is the same as defined in the first column of Equation (A6) and $Z$ is now the "within-group" variance of the response minus twice $\hat{\alpha}^2(alt)$. In this case, however, use the robust linear regression procedure <u>without</u> an intercept. The resulting estimated slope ($\hat{\beta}_1$) provides the value for $\hat{\sigma}_\delta^2(alt)$.

## 3.2 Case 2: $\hat{\sigma}_\delta^2 < 0$

In this case, the hypothesized random cell-specific proportional effect is not detectable. An alternative estimate for $\alpha^2$ can be determined with Equation (A24), where $\left(Y_1, Y_2, \ldots, Y_{n_j}\right)$ are the $n_j$ values of the response (*e.g.*, relative resistance) for the $j^{th}$ of $N$ groups defined by stress condition and time. An alternative estimate for $\hat{\sigma}_\delta^2$ would be to set it equal to zero, $\hat{\sigma}_\delta^2(alt) = 0$.

$$\hat{\alpha}^2(alt) = \frac{1}{2} \cdot \frac{\sum_{j=1}^{N}\left(n_j - 1\right) \cdot Var\left(Y_1, Y_2, \ldots, Y_{n_j}\right)}{\sum_{j=1}^{N}\left(n_j - 1\right)} \tag{A24}$$