

Lithium Battery Analysis: Probability of Failure Assessment Using Logistic Regression

Travis A. Moebes, PhD
Science Applications International Corporation (SAIC)

Abstract

Fourteen-hundred rows by 53 columns of vendor cell acceptance data [1] were processed through Logistic Regression using SAS® Enterprise Miner® (EM) to find any significant correlation between 52 test output parameters (independent variables) and the pass/fail outcome for each of the 1,400 battery cells tested. The goal was to find helpful predictors for detecting “good” or “bad” cells in the form of a best Logistic Regression model. Data from five cells selected by Johnson Space Center’s (JSC’s) Energy Systems Division (ESD) were processed through three model options (Option1, Option2, and Option3) to determine the best model and to indicate a known cell that failed. The output from the best model showed good acceptability statistics and indicated the failed cell was less acceptable than the other cells.

1. Introduction

This assessment was performed to support the ESD in closing the Lithium Ion Battery (LIB) Serial Number (S/N) 1X1X internal short failure investigation (Failure Investigation Anomaly Report (FIAR) #JSCEP0232) and to also support further battery tests. LIB S/N 1X1X is a flight unit that passed all acceptance testing in October 2005. The battery was brought out of storage on February 2, 2006, to support LIB charger power quality testing. During testing, one of the battery’s five cell modules read 10 mV instead of 3.78V. An investigation determined that shorts internal to a cell were the root cause. The LIB S/N 1X1X status was presented to the Engineering Activities Review Board (EARB) on July 19, 2006 [2].

2. Objectives

The objectives of this assessment are:

- Determine a Logistic Regression model to estimate the probability of a cell passing or failing based on vendor data.
- Apply the model to a given module (e.g., FGM-91) to find cells (serial numbers) that are most likely to fail (e.g., 11YYY and 11XXX).
- Estimate the statistical confidence of the models generated in terms of acceptance statistics, such as *t*-statistics, analysis of deviance (R Square), and Likelihood Ratio Chi-Square test parameters for global null hypothesis.
- Build and test all models on SAS EM data mining tools. Further test models using SAIC in-house software to satisfy SAIC testing requirements.

3. Background

This section gives an overview of the LIB Failure Investigation Destruct Physical Analysis (DPA) results. The hardware descriptions were, in part, extracted from the “Lithium Ion Battery (LIB) System GFE for the Spacesuit” presentation [2]. This section also provides information about the scope, assumptions, and data sources used for this probability of pass/fail assessment.

3.1. Cells 11XXX (cell 5) and 11YYY (cell 4) Cu tab side ground zero cell damage

ESD tests showed there were two separate internal shorts that occurred in cell 11XXX, both on bottom corners of the cell. The short at Ground Zero (GZ) affected several cell fold layers more deeply than the more superficial damage caused by the short at the opposite corner on cell 11XXX. The damage in cell 11YYY could have been collateral to what occurred in cell 11XXX. A one-minute/axis random vibration acceptance test was performed with the LIBs on an open circuit at 30% System-on-Chip (SoC). Open Circuit Voltage (OCV) tests immediately before and after each axis of vibration recorded no changes $> \pm 5$ mV temporal bounding volume (tbv). Nine (tbv) days later, the battery was charged, discharged, and recharged to 30% SoC within family individual cell modules voltages. The pre-test capacity cycling test was performed 13 (tbv) days prior to the vibration testing, and its results were consistent with the post-test capacity cycling results. If two internal shorts developed during the vibration, they must have been of high enough impedance to escape detection during the capacity cycling, and then, transitioned to low impedance shorts to cause the thermal damage on the adjacent corners of cells 11YYY and 11XXX. ESD asked Science Applications International Corporation's (SAIC's) Safety and Mission Assurance (S&MA) Reliability and Maintenance (R&M) personnel to evaluate pass/fail vendor data to see if the vendor data would support predictive analysis for battery failures, even when the vendor data indicated no failures.

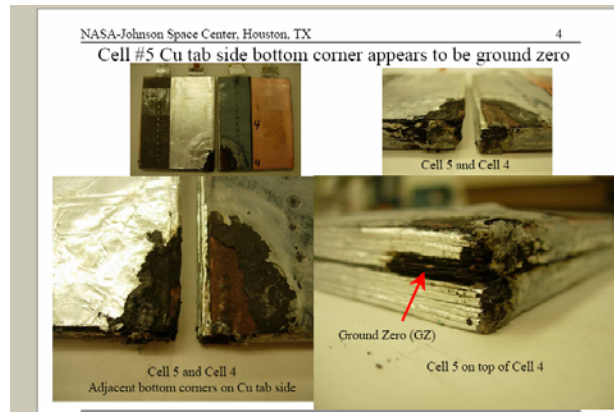


Figure 1. Cells 11XXX (cell 5) and 11YYY (cell 4) damage due to debris

3.2. Scope, assumptions, and data sources

Fourteen-hundred rows by 53 columns of vendor cell acceptance data were obtained from ESD [1]. A portion of the data appears below in Figure 2.

	A	B	C	D	E	F	G	H	I	J
1	LIB No.	Module No.	Serial	P/F	Unfilled cell	Filled cell	Pre-Evac	Post-Evac	Remaining	Mass before dryout, g
2	#N/A	#N/A	10105	Fail	144.5895	160.8774	160.8079	160.1069	15.5869	157.3062
3	#N/A	#N/A	10110	Fail	143.9621	160.0885	160.0064	159.9587	16.0587	157.3732
4	#N/A	FGM-15	10221	Fail	146.9520	163.0626	163.0185	162.7387	15.8308	160.1390
5	#N/A	#N/A	10103	Fail	144.6959	160.8963	160.8312	160.0992	15.4684	157.3956
6	#N/A	FGM-17	10206	Fail	147.0713	163.1838	163.1443	162.8080	15.7762	159.9786
7	#N/A	#N/A	10082	Fail	146.2629	161.9447	161.8704	161.5316	15.3430	158.8710
8	#N/A	#N/A	10085	Fail	144.2636	160.5167	160.4580	160.1366	15.9317	157.2726
9	#N/A	#N/A	10189	Fail	147.2255	163.3931	163.3386	163.2465	16.0755	160.5955
10	#N/A	#N/A	10099	Fail	144.5800	160.5017	160.4289	160.3780	15.8708	157.7142
11	#N/A	#N/A	10093	Fail	145.6345	161.6568	161.5946	161.0680	15.4957	158.3041
12	#N/A	FGM-20	10060	Fail	143.7608	160.0248	159.9527	159.9113	16.2226	157.1666
13	#N/A	#N/A	10106	Fail	144.1928	160.4007	160.3393	159.5444	15.4130	156.8334
14	#N/A	#N/A	10097	Fail	144.4535	160.6940	160.6323	160.0837	15.6919	157.4064
15	#N/A	FGM-18	10252	Fail	144.4783	160.6060	160.5612	160.0363	15.6028	157.3471
16	#N/A	#N/A	11582	Fail	146.1838	162.4224	162.3861	161.9366	15.7891	159.2901
17	#N/A	#N/A	10610	Fail	140.9904	157.1673	157.1350	156.6698	15.7117	153.8728

Figure 2. Snapshot of the vendor cell acceptance data

The LIB No. column refers to the battery, the Module No. column refers to the modules containing the cells, and the Serial column refers to the cells. The P/F column indicates if a particular cell (serial number) passed or failed the vendor test. The parameters in the remaining columns indicate battery characteristics such as charge level and mass before and after tests with both categorical and numerical parameters. The SAIC analyst assumes the

possibility of building a Logistic Regression classification model with vendor data, using pass/fail (P/F) as the dependent parameter and all other parameters as possible independent parameters. The scope of this assessment was limited to the Logistic Regression classification for model determination. Other models such as Classification Trees and Classification Neural Networks were not used. This assessment was also limited to the vendor cell acceptance data set. Different data sets and models are planned for future studies.

4. Analysis and assessment methodology

Since this analysis assumes pass/fail criteria from the model data with both continuous and categorical parameters, the analyst chose Logistic Regression to build the models. All models were built with SAS EM.

4.1. Logistic regression analysis

Logistic Regression is a predictive analysis, like Linear Regression, but Logistic Regression involves predicting a dichotomous dependent variable. The predictors in a regression analysis can be continuous or dichotomous, but Ordinary Least Squares (OLS) Regression is not appropriate if the outcome is dichotomous. The OLS Regression uses a normal probability theory; whereas, Logistic Regression uses a binomial probability theory. The binomial probability theory makes this analysis a bit more complicated mathematically, so it is only covered superficially in this document.

4.1.1. Chi-square and logistic regression. Since the Logistic Regression and Chi-Square analyses both use the binomial distribution, one might expect there to be a relationship between the two. It turns out that the 2 X 2 contingency analysis with Chi-Square is really a special case of Logistic Regression, and this relationship is analogous to the relationship between Analysis of Variance (ANOVA) and regression. With Chi-Square contingency analysis, the independent variable is dichotomous and the dependent variable is dichotomous. An equivalent Logistic Regression analysis with a dichotomous independent variable predicting a dichotomous dependent variable can also be conducted. Logistic Regression is a more general analysis; however, the independent variable (i.e., the predictor) is not restricted to a dichotomous variable. Logistic Regression is also not limited to a single predictor.

For example, Coronary Heart Disease (CHD) is an increasing risk as one's age increases. Therefore, think of CHD as a dichotomous variable (although one can also imagine some continuous measures of this). For this example, a patient either has or does not have CHD. Figure 3 is a scatter plot that shows the relationship between age and CHD.

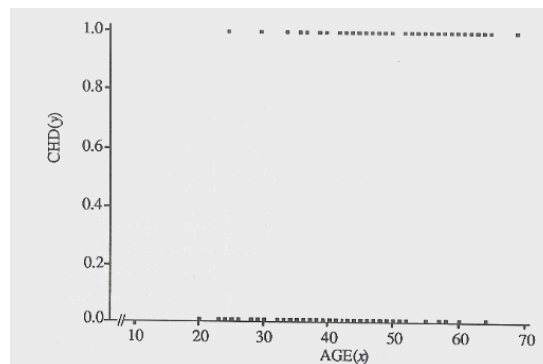


Figure 3. Plot of CHD by age

The scatter plot above shows there is somewhat of a greater likelihood that CHD will occur at older ages; however, this figure is not suitable for examining this relationship. Drawing a straight (best fitting) line through the points also does not do a sufficient job of explaining the data. One solution is to convert or transform the numbers into probabilities. For example, the analyst might compute the average of the y values at each point on the x axis. The y values can only be 0 or 1, so an average of them will be between 0 and 1 (.2, .9, .6, etc.). This average is the same as the probability of having a value of 1 on the y variable, given a certain value of x (notated as $P(y|x)$). Then,

the probabilities of y at each value of x could be plotted in a plot diagram. Figure 4 shows the scores for CHD and age transformed to probabilities in a plot diagram.

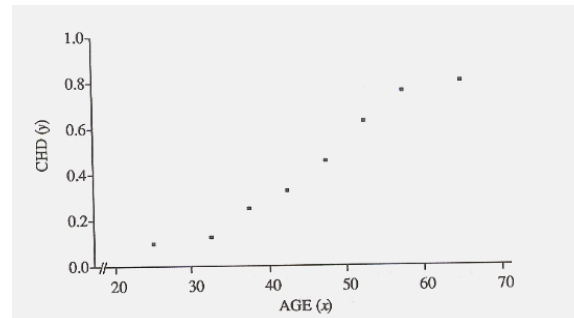


Figure 4. Cumulative probability curve for the logistic distribution

The transformed probabilities produce a smoother curve so that the curve now fits a cumulative probability curve for the logistic distribution, and it is easy to see that the probability of having CHD increases as the values of x increases. The curve in Figure 4 is not a straight line; it is more of an s-shaped curve. However, this s-shape resembles some statistical distributions that can be used to generate a type of regression equation and its statistical tests.

4.1.2. The logistic regression equation. Further mathematical transformations are needed to get from the s-curve (as in logistic) shown in Figure 4 to a straight line (as in regression). The formula below has a natural logarithm in it:

$$\ln \left[\frac{p}{1-p} \right] = \alpha + \beta x \quad \text{Eq. 1}$$

Equation 1 shows the relationship between the regression expression (on the right), which is a straight line expression, and the Logistic Regression expression (on the left). Equation 1 involves the probability, p , that y equals 1 and the natural logarithm (a mathematical function abbreviated $\ln$). In the CHD example, the probability that y equals 1 is the probability of having CHD. Equation 2 is used to determine the probability of a person having CHD based on age.

$$p = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)} \quad \text{Eq. 2}$$

The above formula, called the logit transformation, uses an abbreviation for exponent ($\exp$). $\exp$, the exponential function, involves a constant with a value of 2.71828182845904 (roughly 2.72). To take the exponential function of a number, raise 2.72 to the power of the number. For example, $\exp(3)$ equals 2.72 cubed or $(2.72)^3$, which equals 20.09. The natural logarithm is the opposite of the $\exp$ function. The $\ln(20.09)$ is the number 3. $\exp$ and $\ln$ are common mathematical functions on many calculators.

Although a software tool (e.g. SAS EM) perform the calculations for Eqs. 1 and 2, they are provided in this paper to demonstrate how to get from a regression formula for a line to the logistic analysis and back. Logistic Regression analysis follows a very similar procedure to OLS Regression. The only difference is the need to transform the regression formula and some binomial theory to conduct the tests.

4.1.3. Model fit and the likelihood function. Similar to regression, one can find the best fitting line of sorts for the Logistic Regression curve. In regression, analysts use a criterion called OLS, which minimizes the squared residuals or errors to find the line that best predicts the swarm of points. In Logistic Regression, analysts use a slightly different system. Instead of minimizing the error terms with least squares, analysts use a calculus-based function called Maximum Likelihood (or ML). In Logistic Regression, ML finds the function that will maximize

the analysts' ability to predict the probability of y based on their knowledge about x . In other words, ML finds the best values for the formulas discussed above (Eqs. 1 and 2) to predict CHD with age.

The ML function in Logistic Regression provides a kind of Chi-Square value. The Chi-Square value is based on the ability to predict y values with and without x . This is similar to what was done in regression in some ways. Remember that the ability to predict y is based on the distance between the regression line and the mean (the flat, horizontal line) of y . The Sum of Squares Regression is based on the difference between the predicted y and the mean of y ($SSR = \sum (\hat{y} - \bar{y})^2$). Another way of stating this is that regression analysis compares the prediction of y values when x is used to predict them. The ML method performs a similar function. It calculates the fitting function without using the predictor, x , and then recalculates it using what is known about x . The result is a difference in goodness-of-fit. The fit should increase with the addition of the predictor variable, x . Thus, a Chi-Square value is computed by comparing these two models (one using x and one not using x).

The conceptual formula is shown in Eq. 3, where G stands for "goodness-of-fit":

$$G = \frac{\text{likelihood with the predictor}}{\text{likelihood without predictor}} \quad \text{Eq. 3}$$

Mathematically speaking, Eq. 3 is more precisely described in Eq. 4.

$$G = -2 \ln \left[\frac{\text{likelihood without the predictor}}{\text{likelihood with the predictor}} \right] \quad \text{Eq. 4}$$

Sometimes G is referred to as "-2 log likelihood," as SAS does. G is distributed as a Chi-Square statistic with 1 degree of freedom, so a Chi-Square test is the test for the fit of the model. As it turns out, G is not exactly equal to the Pearson Chi-Square, but it usually leads to the same conclusion.

4.1.4. Odds ratio and slope (β) values. Slope values, β , can be obtained from the logistic equation (Eq. 2). These values are a result of transforming equations that allowed us to get from the logistic equation to the regression equation. The slope can be interpreted as the change in the average value of y , from one unit of change in x .

The Odds Ratio (OR) is also obtained from the Logistic Regression. The OR is equal to the exponential function of the slope, calculated as $\exp(\beta)$. The OR is interpreted as it would be with a contingency table analysis. An OR of 3.03 indicates that there is about a three-fold greater chance of having CHD given one unit increase in x (e.g., 1 year increase in age). If this was the ratio obtained from the age and CHD example, the OR would indicate a 3.03 times greater chance of having CHD with every year increase in age.

Converting from the slope, β , to the OR is relatively easy with most calculators, as shown below.

$$\begin{aligned} OR &= \exp(\beta) \\ \beta &= \ln(OR) \end{aligned}$$

4.1.5. Coding Y and X. It is important to code the dichotomous variables as 0 and 1 with Logistic Regression (rather than 1 and 2 or some other coding), because the coding affects the ORs and slope estimates.

Also, if the independent variable is dichotomous, SAS EM asks the user's coding method preferences for the predictor variable (under the Categorical button), and there are several options to choose from (e.g., difference, Helmert, deviation, simple). All the choices are different methods of dummy coding that will produce slightly different results. The recommended coding scheme for most situations is "indicator." Usually, the absence of the risk factor is coded as 0, and the presence of the risk factor is coded 1.

4.2. Logistic regression model assessment methodology

4.2.1. SAIC Excel Application. An SAIC Excel Application (SEA) uses coefficient estimates based on t -statistics and analysis of deviance (R Square) methods for Logistic Regression Model Assessments. Figure 5 is a sample output from the Logistic Regression application in SEA. The output is displayed in three tables of statistics. These tables include the following information:

- Coefficients for the model, their corresponding standard errors, and t -statistics
- Analysis of deviance statistics
- Importance statistics, showing the significance of each term in the model

The following subsections briefly explain the statistics and statistical parameters included in the SEA output tables. For more detailed information about Logistic Regression in SEA, refer to *Statistics For Managers* [3] and the *Generalized Linear Models, 2nd Edition* [4].

Logistic Regression				
DEPENDENT VARIABLE: KYPHOSIS				
Coefficient Estimates				
Variable	Estimate	Std.Err	t-Statistic	Pr(t)
(Intercept)	-2.04	1.45	-1.41	0.16
Age	0.01	0.01	1.70	0.09
Number	0.41	0.22	1.83	0.07
Start	-0.21	0.07	-3.05	3.1E-3

Analysis of Deviance		
Source	DF	Deviance
Regression	3	21.85
Error	77	61.38
Null	80	83.23

Term Importance			
Source	Wald Statistic	DF	Pr
Start	9.30	1	2.3E-3
Number	3.33	1	0.07
Age	2.87	1	0.09

Figure 5. SEA display for the logistic regression node

4.2.1.1. Coefficient estimate parameters. A t -statistic measures the significance of a variable in the model by determining whether the coefficient is significantly different from zero. A t -statistic is the resulting ratio when the coefficient estimate is divided by its standard error. In general, a t -statistic greater than 1.96 in magnitude indicates the coefficient is significantly different from zero, and the associated variable should therefore be kept in the model. The t -statistics in Figure 5 imply the coefficients for **Start** and **Intercept** are significant in the model.

The p -value for each t -statistic indicates if the corresponding coefficient is significant in the model. In general, a p -value less than 0.05 and a t -statistic greater than 1.96 suggests that a coefficient is significant. The small p -value ($\text{Pr}(|t|)$) for **Start** in Figure 5 implies the term is significant, and the variables **Age** and **Number** contribute less to the model. Testing for the intercept is generally not necessary since the regression surface is expected to rarely intersect with the origin.

4.2.1.2. Analysis of deviance. The Analysis of Deviance table includes the regression, error, and null deviances as well as the corresponding degrees of freedom (DF). Deviance is a measure of model discrepancy used in Logistic Regression. Similar to the analysis of variance for a linear model with Gaussian errors, the analysis of deviance is a partition of a model discrepancy from an oversimplified model (i.e., the null model, which consists of two components — regression and error). The null deviance is analogous to the total sum of squares in an analysis of variance. The regression deviance measures the portion of the null deviance explained by the model. The error is what is left over from a perfect model fit. Further refining the regression deviance includes partitioning it by each independent variable; however, this is impractical in data mining since it would require dropping each variable, refitting the model, and taking the difference in the error deviance.

R Square is the coefficient of determination. It is calculated by dividing the regression deviance into the null deviance. Based on the values in the Analysis of Deviance table in Figure 5, R Square is 0.2625.

$$21.85 / 83.23 = 0.2625$$

This R Square value indicates that 26.25% of the variation in the dependent variable is explained by the variation in the independent variables. One should expect at least 50%. The higher the R Square, the better the model. Therefore, one should generally choose the model with the largest R Square statistic when comparing multiple models.

4.2.1.3. Term Importance. The Term Importance statistic can give a substitute for categorical, independent variable assessment in the refined SEA model.

4.2.2. SAS Enterprise Miner. SAS Enterprise Miner uses the Likelihood Ratio Test for Global Null Hypothesis to evaluate the Logistic Regression model, as shown in the output in Figure 6. The Null Hypothesis is that there are no significant variables in the model whose variation explains the variation in the dependent variable. One may reject the Null Hypothesis if Pr is less than 0.0005 (95% confidence) and Beta equals 0. Models with higher Likelihood Ratio Chi-Square statistics are the better models.

Likelihood Ratio Test for Global Null Hypothesis: BETA=0				
-2 Log Likelihood Intercept Only	Likelihood Intercept & Covariates	Likelihood Ratio Chi-Square	DF	Pr > ChiSq
1608.285	820.728	787.5574	38	<.0001

Figure 6. SAS EM display for the Logistic Regression node

For more information about Logistic Regression in SAS EM, refer to the *Predictive Modeling Using SAS Enterprise Miner 5.1* [5] and the *Generalized Linear Models, 2nd Edition* [4].

5. Analysis and assessment results

Data from five cells selected by JSC's ESD were processed through three model options (Option1, Option2, and Option3) to determine the best model and to indicate a known cell that failed. The output from the best model showed good acceptability statistics and indicated the failed cell was less acceptable than the other cells. The processing and model building results were similar in both SEA and SAS EM.

5.1. Lithium ion battery analysis options

After several Logistic Regression analysis iterations, three options were selected as the best candidates to forecast the pass/fail probability of a LIB cell, based on the results of the vendor's cell acceptance data. The following subsections explain the three modeling options.

5.1.1. Option 1. A total of 1,400 rows of data and all usable columns (53) were used as input to initiate the process for finding the best model. There were 52 independent variables and 1 dependent variable.

5.1.2. Option 2. This option is like Option 1, except six independent variables were dropped, because they had less significant statistical indications based on *p*-values and *t*-statistics from Option 1.

5.1.3. Option 3. This option is also like Option 1, except the following columns were deleted:

- Columns with too many blanks, namely the "soft short test parameters" with the exception of column ACR8C.
- Non-relevant columns according to JSC's ESD. Relevant columns with good *t*-statistics were kept.
- Categorical columns, because they had less than desirable *t*-statistics.

5.2. Lithium ion battery analysis results

The following table summarizes the model evaluation analysis results from both SEA and SAS EM.

Table 1. Model evaluation analysis results

	R-Square	LR Ratio	Pr/Beta
Option 1	0.0750	552.37	0.0001/0.0
Option 2	0.6190	223.49	0.0001/0.0
Option 3	0.6900	787.55	0.0001/0.0
	SEA	SAS EM	SAS EM

R Square calculations are obtained from SEA. Likelihood Ratio Chi-Square (LR Ratio) and Pr/Beta calculations are obtained from SAS EM. Both SEA (R Square) and SAS EM (LR Ratio) indicate Option 3 produces the best model.

The following tables summarize the Pass/Fail output from both SEA and SAS EM

Table 2. Pass/fail output from SEA and option 3

PREDICT.prob	PREDICT.class	Cells
0.794243293	Pass	11XXX
0.79909671	Pass	11YYY
0.982199879	Pass	11AAA
0.921305376	Pass	11BBB
0.900000022	Pass	11CCC

Table 3. Pass/fail output from SAS EM and option 3

PREDICT.prob	PREDICT.class	Cells
0.792776	Pass	11XXX
0.799096	Pass	11YYY
0.982199	Pass	11AAA
0.921305	Pass	11BBB
0.900899	Pass	11CCC

The PREDICT.prob column in the tables above is the model's prediction that the cell should pass vendor tests. The PREDICT.class column is the model's indication of pass or fail.

Based on the PREDICT.prob values in Tables 2 and 3, model Option 3 in both the SEA and SAS EM applications indicate that cells 11XXX and 11YYY should pass, but with a predictive probability significantly less (or at least noticeable) than the other cells. Cell 11XXX definitely failed during JSC ESD testing, and cell 11YYY may have failed or may have suffered collateral damage.

6. Conclusions

Conclusion 1: Modeling Option 3, the best statistical model, showed that cells 11XXX and 11YYY in module FGM-91 and battery 1X1X had a probability of 0.79 to pass. This was the lowest probability of passing than any other cells in the module. All other cells showed a 0.90 probability or higher to pass. Cell 11XXX failed during a critical NASA exercise, and cell 11YYY showed problems during ESD tests. Inspections showed debris in the problem cells. Acceptance statistics for modeling Option 3 were good. This model could be used to test all battery data and assist management in further safety testing.

Conclusion 2: Model Option 3 was initially constructed with all 1,400 rows and 53 columns of vendor cell acceptance data. The final version of the model was built with parameters that were designated as relevant by JSC's

ESD and that had good t -statistics in the initial model. With some fine-tuning, Option 3 should be a good approach for building a model with battery data.

Conclusion 3: The SEA and SAS EM features proved to be helpful in this analysis. Since both applications suggested the same modeling option as the best model, there is a better confidence in the final model and its results.

7. Recommendation

Based on the results of this analysis, Option 3 is the recommended model for analyzing battery data. The Logistic Regression processing in Option 3 produced good, predictive analytics. If required, a complete Logistic Regression analysis or other classification analysis of all vendor battery data could be performed to assist testing management.

8. References

- [1] Cell Acceptance Data Spreadsheet 1, Electroveya, August 10, 2006, Original file name: EV307717D_spreadsheet.xls.
- [2] Darcy, Eric et al, "Lithium Ion Battery (LIB) System GFE for the Spacesuit," Power Point Presentation, August 11, 2006.
- [3] Levine, David et al, "Statistics For Managers Using Microsoft Excel, Fourth Edition, Prentice Hall, Upper Saddle River, New Jersey, June, 2004.
- [4] McCullagh, P. and Nelder, J.A., Generalized Linear Models (2nd ed.), Chapman and Hall, London, 1989.
- [5] *Predictive Modeling Using SAS Enterprise Miner 5.1*, Book Code 59927, SAS Institute Incorporated, Irvine, California, September, 2004.

8. Contact Information

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

Travis A. Moebes, PhD
SAIC
2450 NASA Parkway (NASA Road One) #224D
Houston, Texas 77058
Work Phone: 281-335-2288
Fax: 281-335-3482
E-mail: travis.a.moebes@nasa.gov
Web: www.saic.com

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