



Performance Degradation Analysis for Reliability/Durability Testing

1 Introduction

Note: Nothing in this standard supercedes applicable laws and regulations.

Note: In the event of conflict between the English and domestic language, the English language shall take precedence.

1.1 Purpose. This standard provides instructions and guidelines for implementing Degradation Analysis, as part of a Weibull analysis, to determine product reliability. In Degradation Analysis, the performance or condition of each part is measured repeatedly throughout the test, with the time to failure (defined as the performance/condition degrading to some specified allowable limit of acceptability) being predicted for any parts that have not failed by the end of testing. These **predicted** failure times are included along with any **actual** failure times in the Weibull analysis. Using these predicted times to failure provides more information to the Weibull analysis than does calling these unfailed parts right censored (or suspended) at the point they were taken off test.

Degradation Analysis may be used for both Development and Validation of a product. It is primarily used in a Test-to-Failure validation plan, where, in effect, it increases the number of failure points, and thus improves the Weibull analysis. Both time and cost can be minimized while still implementing a Test-to-Failure plan, because prediction of future failure times reduces the number of actual failures needed to get a good Weibull analysis and reliability assessment. Degradation Analysis may also be used for a Success Test plan, as the measured performance and predicted lives add valuable information about the product and its capabilities, as well as giving an early warning of impending failures.

1.2 Applicability. This standard can be applied to performance metrics that are used as part of a definition of failure in a product reliability requirement. Such failures are often called “soft failures” because there is still some functionality, even though it is at a level that is not considered acceptable. The identification and setting of customer-focused performance metrics is one of the steps in the Design for Six Sigma (DFSS) methodology.

It can also be used with measurements of a part’s condition, such as the thickness of the brake lining, which is not a measure of performance directly, but which affects performance when the condition changes enough. For simplicity in the remainder of this specification, both direct measurement of performance, and measurement of a condition that affects performance, will be referred to as “performance.”

The following conditions determine its applicability:

- a. The performance metric must indicate some aspect of the product’s operation, with a definable point separating acceptable from unacceptable performance.
- b. The performance must change in a continuous manner during the test, as opposed to a long period with no change followed by a rapid change at the end. (Note that “symptoms” such as noise typically do not meet this condition, so they do not usually make good candidates for Degradation Analysis. Besides not having a set of performance data to extrapolation from, there is also little time savings because the failure occurs soon after the start of the symptoms.)
- c. The test content must be able to elicit this performance change.
- d. The performance must be objectively measureable during the test. (Note that a subjective rating is **not** an appropriate measure for Degradation Analysis – besides not being objective, the rating scale is generally not linear.)

Degradation Analysis is primarily applied to components and subsystems being tested in a lab durability test, although it may also be applied to a vehicle-level test (e.g., tire tread depth on a vehicle durability test).

Degradation Analysis can be used for both “normal stress” testing, and for Calibrated Accelerated Life Testing (CALT) where different parts are tested at one of at least three different overstress levels, with the lives at the “normal” stress being predicted using a life vs. stress curve.

2 References

Note: Only the latest approved standards are applicable unless otherwise specified.

2.1 External Standards/Specifications.

None

2.2 GM Standards/Specifications.

GMW8758

GMW15760

2.3 Additional References.

- Applicable Component or Subsystem Technical Specification (CTS or SSTs).
- Applicable Component or Subsystem Test Procedure.
- Help function in Weibull++ or ALTA software.
- ReliaSoft HotWire newsletters, Issues 86 and 149, located at:
<http://www.weibull.com/hotwire/archive.htm>.
- User Guide to the Weibull++ or ALTA software, available through the program’s File/Help tabs, or the combined Weibull++/ALTA guide is currently at:
http://www.synthesisplatform.net/WeibullALTA/en/UG_WeibullALTA9.pdf.

3 Resources

3.1 Facilities. The test facilities shall be as specified in the applicable product requirement and/or test procedure documents.

3.1.1 Calibration. The test facilities and equipment shall be in good working order and shall have a valid calibration label.

3.1.2 Alternatives. Alternative test facilities and equipment may also be used. However, all measuring variables as specified in this standard shall be determined correctly with respect to their physical definition.

3.2 Equipment. The test equipment shall be as specified in the applicable product requirement and/or test procedure documents.

3.3 Test Vehicle/Test Piece. The test vehicles and/or test pieces shall be as specified in the applicable product requirement and/or test procedure documents.

3.4 Test Time. Not applicable.

3.5 Test Required Information. Not applicable.

3.6 Personnel/Skills. Not applicable.

3.7 Software. Degradation Analysis requires the use of software that can:

- a. Determine the one best type of curve (linear, exponential, etc.) to fit to all the performance data for all test pieces.
- b. Perform the curve fit for each part being tested.
- c. Extrapolate each curve to the level of performance that defines the failure, to get predicted times-to-failure. This capability may be integrated with the Weibull analysis software (as in ReliaSoft’s Weibull++ and ALTA, which is available within GM), or it may be stand-alone software that requires the user to transfer the life data to a separate Weibull analysis package.

4 Procedure

4.1 Preparation.

4.1.1 Develop Requirements.

4.1.1.1 Identify products that have at least one performance measure that may allow the use of Degradation Analysis. See Appendix A for a list of “thought starters” – parts and performance measures that might be good candidates for Degradation Analysis.

© Copyright 2014 General Motors Company All Rights Reserved

September 2014

Page 2 of 24

4.1.1.2 Verify that the performance measure meets the conditions listed in 1.2 Applicability.

4.1.1.3 Verify the feasibility of measuring the performance during the test. This includes verifying the availability of measurement equipment and the ability to take the measurements on a regular basis.

4.1.1.4 Define the critical performance level that defines a failure.

4.1.1.5 Determine the required measurement and frequency. This may be continuous automated monitoring, or periodic measurements that occur based on calendar time (e.g., daily) or test time (based on test hours, cycles or percent of test). There needs to be enough points in the full curve to adequately define its curvature, if any. The shorter the test, that is, the quicker failures can occur, the more often the measurements need to be made. A recommended frequency is at least every 20% of a test life, assuming failures are expected to occur after one life. This helps protect against having too few data points if the performance degrades faster than expected.

4.1.1.6 Determine the required sample size. The greater the number of parts, the more likely it is that the most appropriate curve fit will be found. In addition, the subsequent Weibull analysis is greatly improved with higher sample sizes. When using Degradation Analysis, it is better to have more parts and test them a shorter amount of time, rather than the opposite, since the extrapolations give more "failures" without actually testing that long. A sample size of at least six is recommended.

4.1.1.7 Make any necessary changes to the applicable test procedure to specify the monitoring requirements.

4.1.1.8 Make any necessary changes to the applicable product requirement document (e.g., CTS or SSTs) to specify the product's performance requirements.

4.1.1.8.1 This typically will include identifying the performance metric, stating the minimum (and/or maximum, if a higher level is worse) acceptable performance level, and specifying any minimum sample size and test length requirements. See Appendix B for example wording.

4.1.1.8.2 Note that this is in addition to other aspects of the reliability requirement, namely: the reliability (R) and confidence (C) numbers (e.g., $R \geq 99\%$, $C = 50\%$); the definition of "one life" on test (given in terms of the number of cycles or hours of some type of operation); other aspects of the usage and environmental conditions (e.g., force, temperature, vibration level); and any other applicable definitions of failure (e.g., hard failures, meaning complete loss of function, and any other soft failures).

4.1.1.9 Related requirements such as part selection (such as GMW15760, which is **required** for selecting Process Validation (PV) and Post Validation Audit (PVA) parts), and data reporting requirements, should also be captured in the sourcing documents.

4.1.2 Data Collection.

4.1.2.1 Measure the performance of each test piece at start of test, or soon thereafter if a "break-in" period is desired (e.g., in a brake test, to give the brake pads time to wear off a surface coating before beginning the measurements).

4.1.2.2 Measure the performance throughout the test at least as often as required by the product specification or test procedure.

4.1.2.3 It is best practice to run long enough for some parts to reach the failure point. This does two things:

- It allows verification that the failure point really is set at an appropriate level, and
- It gives one or more complete curves to help determine the most appropriate curve fit type for extrapolating the remaining unfailed parts.

One to two failures is usually enough for a reasonable Degradation Analysis.

4.1.2.4 The measurement frequency can change during the test (e.g., an increase in frequency near the start of test, to verify the performance is changing, or as a part approaches the failure criteria, to better capture the point of failure).

4.1.2.5 The parts can have their performance measured at different times.

4.1.2.6 For a test where all the parts see the same loads, the data should be recorded as in Table 1.

4.1.2.6.1 The measurement time is when the measurement was taken, relative to the start of test, and can be in hours, cycles, percent of one life, etc.

4.1.2.6.1.1 To be able to calculate all the curve fit types shown in Figure 1 (some of which will not accept a time of zero), any data point at start of test should be listed at a time very close to zero, as illustrated in Table 1.

4.1.2.6.2 The performance measure is in whatever units is appropriate.

4.1.2.6.3 The part identification (“Part ID”) can be any unique code that is appropriate for distinguishing between the test pieces.

Table 1: Performance Data, All Parts Subjected to the Same Loads

Measurement Time	Performance	Part ID
0.01	18.3	1
0.01	17.6	2
0.01	19.1	3
...	...	...
100	18.1	1
100	17.5	2
100	18.8	3
...	...	...

4.1.2.7 For a CALT test (described in GMW8758) where there are several groups of parts, each group being tested at a different load, the data should be recorded as in Table 2.

4.1.2.7.1 At least one “load” column is used to identify the specific test condition applied to each part.

4.1.2.7.2 In this example, three parts each were tested at one of three different loads (temperatures in this case). This is the **minimum** recommended number of parts. This may result in a sample size higher than is typical with a non-overstress test. However, by using higher than normal loads, the time to failure (predicted or actual) is substantially reduced.

4.1.2.7.3 More than one type of load can be varied (such as temperature and humidity). Each additional load type requires one more column in the data table, and an increase in the sample size. This is required for the analysis program to determine how each type of load affects the part life.

Table 2: Performance Data for an Accelerated Test, with Parts Subjected to Different Loads

Inspection Time (Cycles)	Performance Measurement	Temperature (K)	Part ID
0.1	10.4	422.0	1
0.1	11.9	422.0	2
0.1	10.2	422.0	3
0.1	9.7	394.3	4
0.1	10.2	394.3	5
0.1	9.9	394.3	6
0.1	9.0	366.5	7
0.1	10.0	366.5	8
0.1	10.5	366.5	9
10	8.2	422.0	1
10	9.7	422.0	2
10	7.9	422.0	3
10	8.0	394.3	4
10	9.3	394.3	5

© Copyright 2014 General Motors Company All Rights Reserved

Inspection Time (Cycles)	Performance Measurement	Temperature (K)	Part ID
10	8.9	394.3	6
10	8.5	366.5	7
10	9.0	366.5	8
10	9.7	366.5	9
...	...	...	...

4.2 Conditions. Not applicable.

4.2.1 Environmental Conditions. Not applicable.

4.2.2 Test Conditions. Deviations from the requirements of this standard shall have been agreed upon. Such requirements shall be specified on component drawings, test certificates, reports, etc.

4.3 Instructions. Not applicable.

5 Data

5.1 Calculations. The following analysis procedures should be performed at periodic intervals while the test is underway. This helps to verify that Degradation Analysis will be useful for the performance data being collected, as well as providing an “early warning” of impending failures.

5.1.1 Data Entry. The following performance-related information is to be included in the analysis:

5.1.1.1 The data, as illustrated in either Table 1 or Table 2, depending on how the test was performed.

5.1.1.1.1 If there is any data at Time = 0 (at start of test), change each time from zero to a very small number near zero, since some of the degradation models do not allow zero for a time.

5.1.1.2 The level of performance defined as the failure point. This is called “Critical Degradation” in Weibull++ and ALTA. This can be either an upper limit (failure is when the measure goes too high) or a lower limit (failure is when the measure goes too low). Only one of these is allowed in a given analysis.

5.1.1.3 A point at which to stop extrapolating the parts’ performance if it has not reached the failure point by that time. This is called “Suspend After” in Weibull++/ALTA. This should be in the same units as the “inspection time.” More on this entry in 5.1.2 Degradation Curve Fitting.

5.1.2 Degradation Curve Fitting.

5.1.2.1 There are many possible curve shapes. Figure 1 shows the types available in Weibull++ and ALTA:

$$\begin{aligned}
 \text{Linear: } & y = a * x + b \\
 \text{Exponential: } & y = b * e^{(a*x)} \\
 \text{Power: } & y = b * x^a \\
 \text{Logarithmic: } & y = a * \ln(x) + b \\
 \text{Gompertz: } & y = a * b^{c^x} \\
 \text{Lloyd - Lipow: } & y = a - b/x
 \end{aligned}$$

Figure 1: Degradation Curve Types in Weibull++ and ALTA

5.1.2.2 In these equations, y is the performance, x is the time, and a, b and c are parameters calculated to give the best curve fit for each individual part.

5.1.2.3 All the parts’ performance data is fit to the same type of curve, although each part’s curve will have its own unique values for parameters a, b, and c.

5.1.2.4 It is possible to specify one of these shapes, if the form the data should take is known from experience. However, there is limited published guidance for picking any specific type of curve.

5.1.2.5 The more typical way to do the curve fitting is to use the Degradation Model Wizard to let Weibull++/ALTA determine how well each type of curve fits the data. The wizard analyzes each curve type, and recommends the curve that minimizes the error between the parts' data points and the calculated curves. Note, however, that the wizard cannot make an assessment of how reasonable the curve is when extrapolated **past** the data, as is done to predict the times to failure.

5.1.2.6 After running the Degradation Model Wizard, check its results to see if the second (or even third) best curve fit type also fits the data well. Later, it may be appropriate to use a different curve type if the first choice does not give a reasonable extrapolation beyond the last data point. Paragraph C1.3.2 Performance Curve Fitting shows how to assess the acceptability of the various curve types.

5.1.2.7 Three of the curve types, Power, Logarithmic and Lloyd-Lipow, cannot be fit if there is any data given at Time = 0. If there is any such data, simply change the time to a very small number near zero.

5.1.2.8 As mentioned in 5.1.1 Data Entry, the extrapolations can be suspended at some specified time. Any part curves that have not yet reached the failure point will be entered into the Weibull analysis as right censored (suspended) at the point where the extrapolation was stopped. This is used to prevent unrealistic life predictions that may arise from extrapolation on a curve that fit the data well, but may not be appropriate further out. For example, some curves, depending on the parameters, can make a sharp turn at some point. It is good practice to manually increase the upper limit on the x-axis of the degradation plot to see what the curves do further out, and suspend the extrapolation before the curve starts behaving unrealistically.

5.1.2.9 It is recommended to initially suspend the extrapolations at two times the length of actual testing. For example, if testing went to three lives, suspend at six lives. After reviewing the Degradation plot, decide whether or not to use a higher or lower suspension point.

5.1.3 Weibull Analysis.

5.1.3.1 Weibull analysis of the data generated with Degradation Analysis should follow the normal GM guidelines, which generally call for a Rank Regression on X (RRX in Weibull++, also called Least Squares) analysis. There are situations when a Maximum Likelihood Estimation (MLE) analysis is preferred over RRX, but those situations are uncommon when using Degradation Analysis.

5.1.3.2 Only a 2-parameter analysis should be used. (Note that in Weibull++, a 2-parameter analysis where the slope is specified by the user is called a 1-parameter Weibull analysis).

5.1.3.3 Other analysis types could fit the data better than does the Weibull distribution. Weibull++ has a wizard that determines how well other curve types (such as lognormal, generalized gamma, or log-logistic) fit the life data. Use one of the alternatives only if the Weibull distribution gives a noticeably poorer fit.

5.1.3.4 The life-related information in 5.1.3.4.1 thru 5.1.3.4.3 is to be included in the analysis:

5.1.3.4.1 The actual and predicted lives generated from the Degradation Analysis. This data is automatically used in Weibull++, but if the Degradation Analysis and Weibull programs are separate, the failure data will need to be transferred between them manually.

5.1.3.4.2 The type of analysis, per 5.1.3.1 thru 5.1.3.3.

5.1.3.4.3 If using Weibull++, there is a checkbox called "Use extrapolated intervals." If this option is not selected, the Weibull analysis will use the predicted time of each failure as input. If it is selected, the Weibull analysis will instead use a range of values (using a lower and an upper bound) to describe the time to failure. The two choices will generally give different answers, but there is no compelling reason to use the more complicated "intervals" approach (that is, there is no reason why that would give a better answer, despite all the added calculations that need to be made). Furthermore, using the "intervals" option increases the chance that the Weibull analysis will not converge properly. It is, therefore, recommended to not use this option, that is, do **not** check the "Use extrapolated intervals" box.

5.1.4 Special Cases.

5.1.4.1 Multiple Failure Mechanisms.

5.1.4.1.1 Only one type of Degradation Analysis can be performed at a time. However, there may be more than one type of performance being monitored in the testing, each with its own performance measurements and failure criteria. There may also be some failures for which there was no performance data that would allow use of Degradation Analysis.

5.1.4.1.2 In cases where there are multiple types of failures, the question is, "How do you determine an overall product reliability that can be compared to the product's (overall) reliability requirement?"

5.1.4.1.3 If there are enough failures of each type to do a good Weibull analysis, each should be analyzed separately. The overall reliability at one life is the product of the individual failure's reliabilities at one life. That is, $R_{\text{Overall}} = R_1 * R_2 * \dots * R_N$, where the numbered subscripts are for individual failure types, and N is the total number of failure types.

5.1.4.1.4 Here is an example to illustrate how the data is treated when all the parts are being monitored for performance degradation, and there are failures from that degradation as well as failures of a second type. Assume the two failure types can be analyzed separately. The analysis of the second (non-degradation) type of failure will have failure times that are only from those parts that failed in that manner. All the other parts will be right censored (suspended) at the time when they were **actually taken off test (not their extrapolated lives from the Degradation Analysis)**. However, since all the parts were being monitored for performance, Degradation Analysis can predict the life for every part due to degradation, even for the parts that failed from the second type of failure. In other words, use Degradation Analysis for all the parts (suspending the performance extrapolation if appropriate, as described in Section 5.1.2 Degradation Curve Fitting), to get the reliability for the performance degradation failure.

5.1.4.2 Using Degradation Analysis with Accelerated Testing.

5.1.4.2.1 It is possible to use accelerated testing and Degradation Analysis together. There are several ways to do accelerated testing, but the method that is used with Degradation Analysis is referred to, within General Motors, as Calibrated Accelerated Life Testing (CALT). In CALT, some parts are subjected to one load level, some others see a different load level, and still more parts see a third load level. More than three load levels may be used, but that is uncommon. Data entry for combined CALT/Degradation Analysis was illustrated in Table 2.

5.1.4.2.2 All of the load levels in CALT are higher than the non-accelerated, "normal" loads. Part of the analysis is the creation of a damage accumulation model that gives life as a function of load. This model is used to convert the lives at each accelerated load, back to the normal load level (which is usually the load given in the Component or Subsystem Technical Specification). In other words, a third model (life vs. load) is used (besides the degradation model and the Weibull model). This makes it even more important to have reasonably high sample sizes (three parts at each of three load levels is an absolute minimum), so all three models have reasonable curve fits.

5.1.4.2.3 CALT is described in more detail in GMW8758, but note that that specification assumes testing all the parts to failure (not using degradation analysis to predict times to failure). It, therefore, allows for only two parts tested at each load level.

5.1.4.2.4 The program ALTA (for Accelerated Life Testing Analysis), also available from ReliaSoft, can perform all these calculations to generate the normal-stress Weibull curve.

5.1.4.2.5 ALTA has the same wizards as does Weibull++, to help with the curve fits for the degradation data and for the life data (Weibull, exponential and lognormal are the available options). However, the type of curve used for the life vs. load model must be selected by the user based on what load (or combination of loads) is being accelerated. Nine options are available for this model. The program's Help function provides guidance on which one to use, depending on what inputs are being accelerated. See Appendix D for more details on using ALTA.

5.1.4.3 Degradation Analysis When Only One Performance Measurement is Available for Each Part.

5.1.4.3.1 It is possible to perform a Degradation Analysis in situations where only **one** performance measurement per part is possible. There are a couple situations where this may occur:

- a. Doing the performance measurement will damage or destroy the part such that it cannot be returned to test, or
- b. Data comes from field units and it is not feasible to get more than one measurement per unit.

5.1.4.3.2 A large number of parts is needed to make up for the lack of multiple measurements on each part, since there is not only degradation, but part-to-part variation as well. For example, a part with two years of exposure may have better performance than one with only one year, due to better initial performance that could not be measured. A higher sample size is needed to overcome this problem.

5.1.4.3.3 If analyzing test parts using destructive measurements, it is possible to partially offset the need for a larger sample size by reducing the number of times in the test that the performance is measured. For example, instead of measuring performance on a group of parts every 20% of test, do it every 40% of test. It is recommended to measure at least ten parts at each of at least six points in the test, including at start of test. If

the start of test measurement shows a lot of variation between new parts, then an even higher number should be tested.

5.1.4.3.4 If using data from the field, the variation in how the parts have been used as well as in the environmental conditions to which they have been exposed, is an additional factor that can affect the ability to see performance degradation trends. There is little available guidance on required sample sizes for field data. However, because automotive parts of the same age or mileage typically have more than a 10:1 ratio between severe and non-severe usage and exposure, a total sample size of 200 or greater may be necessary.

5.1.4.3.5 Analysis of this data is beyond the scope of this standard, but ReliaSoft has published "HotWire" newsletter articles on two ways to analyze this, one using ALTA (Issue 149, "Using ALTA to Model Destructive Degradation Data") and another using Weibull++ (Issue 86, "Degradation Analysis in Destructive Testing - Part II").

5.2 Interpretation of Results. The act of monitoring and recording the performance data provides valuable information about the product's capabilities. When this performance data is extrapolated to estimate the future failure times in a test-to-failure plan, it has the potential to add significantly to the accuracy of the Weibull analysis, due to the increased number of "failed" parts. However, it is important when reviewing the results to verify the following:

- An appropriate performance measure was specified;
- The performance measurements were taken often enough to be able to capture the shape of the performance vs. time data;
- An appropriate failure criteria was used;
- The performance data does change in a continuous manner over a significant portion of the test time;
- The extrapolations beyond the end of data behave in a reasonable way, or alternatively, that the extrapolations were suspended prior to any unwanted behavior of the curve-fitted lines;
- Weibull analysis best practice as indicated in 5.1.3, were followed.

Provided these factors all appear acceptable, the reliability predictions generated should be of higher quality than if Degradation Analysis was not used and many of the parts were right censored at the end of test.

5.3 Test Documentation. All the performance data and analysis settings used shall be documented, along with all the product information needed to understand the design level, part source, build dates, etc., Weibull++ and ALTA both have fields in their database that can be used for documenting this type of information.

6 Safety

This standard may involve hazardous materials, operations, and equipment. This standard does not propose to address all the safety problems associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.

7 Notes

7.1 Glossary.

Accelerated Life Testing Analysis (ALTA): The name of software from ReliaSoft for analyzing life data generated with accelerated testing.

Accelerated Testing: Increasing the loads in a durability test to reduce testing time.

Calibrated Accelerated Life Testing (CALT): A method of accelerated, overstress testing that allows the prediction of times to failure at the non-overstress loads that are typically defined in the product specifications.

Confidence: Confidence, as used in reliability analysis, is the lower bound on the %Reliability estimate.

Critical Degradation: The level of performance used to define a failure in a Degradation Analysis.

Degradation Analysis: The analysis of performance data collected over time, to predict when the level of performance will reach the Critical Degradation, i.e., the time to failure.

Hard Failure: A product failure in which there is a complete loss of function.

Life Data Analysis: An analysis that fits product life data to a statistical distribution. Weibull analysis is an example.