SANDIA REPORT
SAND2022-13376
Printed September 2022

Sandia National Laboratories

Evaluation of COTS Electronics by Power Spectrum Analysis and Multivariate Data Analysis

Stephanie A. DeJong, Rosalie Multari, Kelsey Wilson, and Paiboon Tangyunyong

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico
87185 and Livermore,
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

ABSTRACT

Power spectrum analysis (PSA) is a fast, non-destructive, sensitive method for examining commercial off-the-shelf (COTS) electronic components. These features make PSA attractive for both component screening and surveillance in support of component reliability efforts. Current analysis methods limit the utility of PSA due to the need to manually examine the results of analysis to identify anomalous parts. This study demonstrates the development and application of a workflow to automate the screening of COTS electronic components. Further, this study demonstrates the use of multivariate algorithms to assess aging of Zener diodes. These workflows can be readily extended to other components, combining the benefits of PSA and multivariate analysis to screen and evaluate COTS electronic components.

ACKNOWLEDGEMENTS

The authors would like to thank the ND-LDRD IAT for support of this work, along with the Aging and Lifetimes program for support of some of the experimental work leveraged in this report. Thanks also to Ariana Beste for discussions and initial modeling work, and the team of technologists involved in the collection of the data used throughout the report. Further, the authors would like to thank folks that supported experimental conventional and PSA aging and testing, including David Canfield, Mary Groves, Lucas Ridgeway, and Brian Molley.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

This page left blank

ACRONYMS AND TERMS

Acronym/Term	Definition
ASIC	application specific integrated circuit
COTS	commercial off-the-shelf
ESD	electrostatic discharge
IV	conventional current-voltage
LV	latent variables
PC	principal component
PCA	principal components analysis
PLS	partial least squares
PLSR	partial least squares regression
PSA	power spectrum analysis
QDA	quadratic discriminants analysis
RMSEC	root-mean-squared error of calibration
RMSECV	root-mean-squared error of cross-validation
RMSEP	root-mean-squared error of prediction
t-SNE	t-Distributed stochastic neighbor embedding

This page left blank

1. INTRODUCTION

1.1. Motivation

Surveillance of nuclear weapons is critically important to the nuclear deterrence mission to provide test data as evidence of the state-of-health of the stockpile. The majority of age-related degradation occurrences are caused by the presence of manufacturing defects (e.g., a contaminant) and/or the existence of localized micro-environments (e.g., humidity) that can be difficult to detect using conventional methods.

Commercial off-the-shelf (COTS) electronic components are utilized throughout the stockpile, and their use requires the assurance of reliability prior to insertion and continued surveillance throughout the components' lifetimes to establish state-of-health. Current methods to evaluate COTS electronic devices include univariate voltage and current methods, and reliability sampling by destructive analysis. Destructive analysis inherently cannot be used as a screening tool, and current analysis methods of conventional data are not sensitive to more subtle variability of interest in these parts, particularly changes related to device age.

Power spectrum analysis (PSA) has been demonstrated to be a powerful technique to detect a range of features in electronic devices, including identifying signs of aging related to underlying physical phenomena and detecting anomalous or counterfeit devices.[1, 2] Current analysis methods limit the utility of PSA data as they are primarily performed by visual inspection of the data. To deploy PSA as part of a large screening effort, the analysis methods and decision making need to be more capable of detecting features/differences that are not obvious to humans conducting visual assessment. Further, better methods for aging analysis need to be developed to enable detection of possible precursors of future failure.

1.2. Power Spectrum Analysis

PSA uses a unique, off-normal biasing scheme to generate distinct power spectra that can be used to detect differences between electronic devices.[2] A test device is normally pulsed with a square-waveform voltage; the device responds to these pulses by loading on the square-waveform voltage. The amount of loading depends on the device's dynamic impedance. The frequency of the square-waveform voltage depends on the device and ranges from 1 kHz to 1 MHz.

The time-domain signal (the square-waveform voltage) is then transformed via Fourier transform to a frequency-domain signal (the power spectrum) that is characteristic of that device. Since PSA is a comparative technique, the power spectrum is further normalized to an appropriate reference spectrum prior to analysis. The primary reason for normalization is to account for any drift in the instrument set-up or other changes in measurement environment over time. In addition, differences between devices are more easily discernable in normalized spectra compared to raw (un-normalized) spectra. [1, 2]

Each PSA spectrum is 1×801 (amplitude vs frequency). Due to the high dimensionality of the data and large amount of correlation within a sample spectrum, these data are well-suited to dimensionality reduction techniques, such as principal components analysis (PCA). For a given data set, PCA decomposes the data matrix X into a set of vectors such that $X = TP^T + E$, where T are scores (analogous to concentrations), P are weights (analogous to spectra), and E is error.[3] The scores in T are then visualized in 2-D or 3-D plots and manually assessed to make in- and out-of-group determinations.

1.3. Goals

The need to visually assess the scores from PCA limits the utility of PSA as a larger scale screening tool in a few ways. First, visually assessing the data can be time-consuming, and the data is limited to only three PCs (principal components). For complex data, three PCs might not be enough to fully describe the data variance to separate features of interest. Second, decision making based on visual inspection could lead to operator bias, which could influence the results. A less biased method uses quantitative metrics to assess similarity so that decisions are made consistently.

The primary goal of this study is to create an approach for 100% screening of COTS electronic devices using PSA in tandem with multivariate analysis tools. Three general approaches were taken in service of this goal.

1. Evaluate the variability of PSA data across sampling events to understand the best practices to develop models of PSA data that are robust over time.
2. Demonstrate an automated method to identify anomalous units from only baseline pristine units that are more readily available.
3. Develop models of PSA and conventional data to estimate COTS device age after accelerated aging, which can be used to support both reliability and surveillance efforts.

Details of the multivariate analysis tools used to achieve these goals are contained in the report below. Recommendations for data workflows to implement larger scale PSA screening of COTS electronics are also discussed.

2. EXPLORING VARIABILITY OF PSA DATA

PSA data are known to demonstrate changes related to instrument variability and measurement environment. To account for this variability, PSA spectra are typically normalized to a reference spectrum. The efficacy of this normalization can be evaluated by examining the PSA spectra collected from samples that are nominally unchanged over several collection events, referred to as control samples. In this section, PSA data of control samples are examined to investigate the potential influence of PSA instrument variability on future model performance, as well as to assess the power of normalization to control variability.

2.1. Sample Structure of ASIC Control Data

PSA data were collected for different lots of application-specific integrated circuit (ASIC) devices manufactured at Microsystems Engineering Science and Applications (MESA) Silicon Fab at Sandia National Laboratories, New Mexico. These data were collected over 32 collection events across four years, from October 2016–June 2020 (see Table 1). During each collection event, ten control samples were measured alongside the samples of interest. These ten samples had the identifiers: 166, 167, 172, 173, 174, 175, 178, 182, 183, and 185. These samples are nominally unchanged during the four years. One control sample (Sample 166) was used as the reference spectrum against which all PSA spectra during the collection were normalized. This device was repeatedly measured during the course of the collection event, after every ~20 spectra. All sample spectra are 801 points in the frequency space. The mean PSA spectrum of each control sample before and after normalization are shown in Figure 1 and Figure 2.

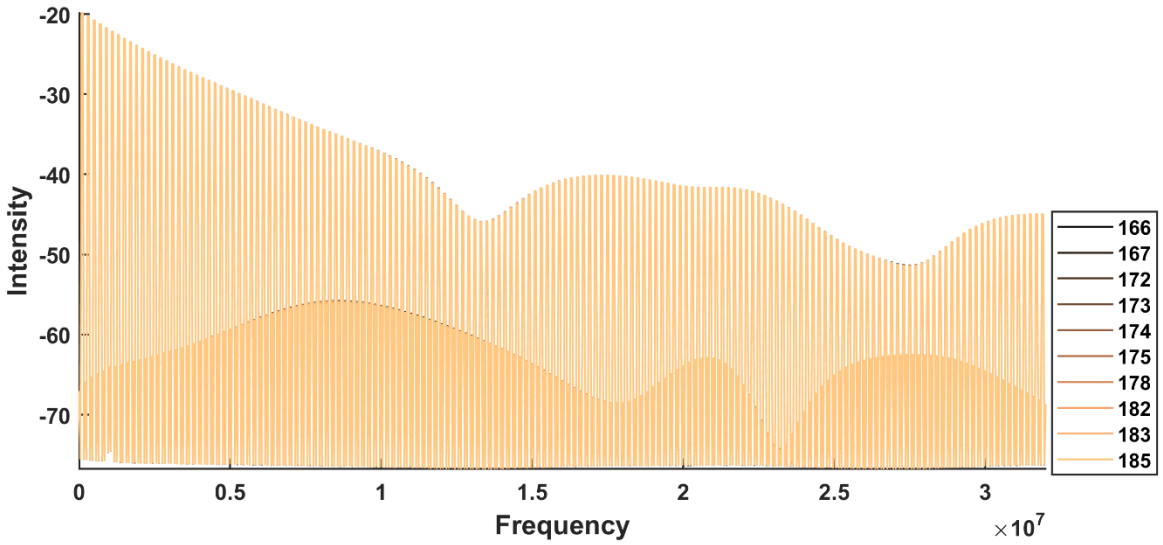


Figure 1. Mean PSA spectra of control samples before normalization to the reference spectrum, collected at 3V (Sample 166)

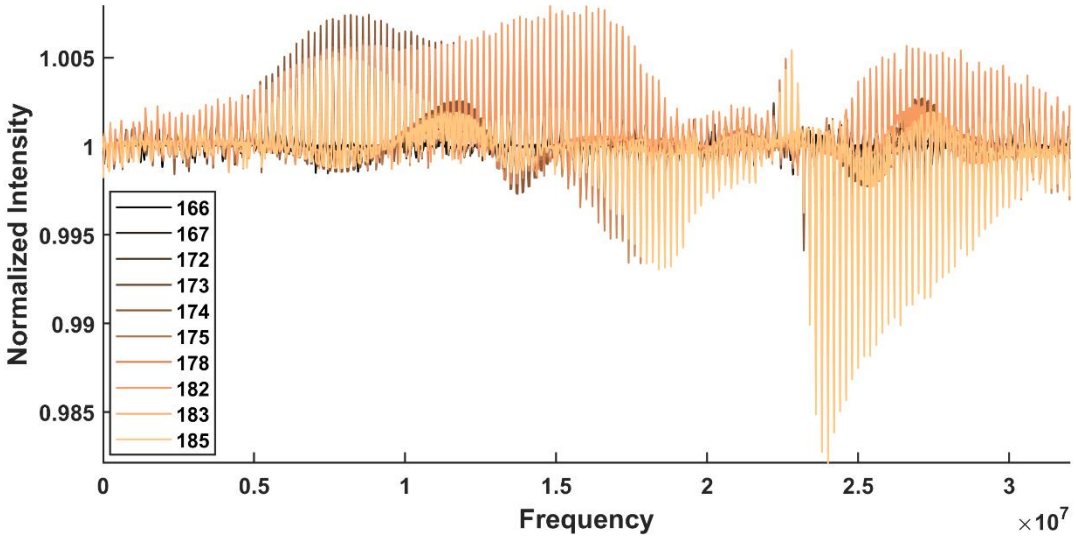


Figure 2. Mean PSA spectra of control samples after normalization to the reference spectrum, collected at 3V (Sample 166)

Table 1. Details of ASIC PSA data

ASIC Lot	Aging Cycles	Collection Date	Reference Replicates	No. Samples
EC Lot	0	14-Feb-2018	6	80
	Radiation	21-Feb-2018	6	80
	Radiation + 1	2-Apr-2018	6	80
	Radiation + 2	18-May-2018	6	80
	Radiation + 3	17-Jul-2018	6	80
Lot 03	0	12-Oct-2016	9	117
	1	4-Apr-2017	9	100
	2	4-Apr-2017	9	99
	3	4-Apr-2017	9	98
Lot 04	0	3-Apr-2017	7	99
	1	4-Apr-2017	7	99
	2	4-Apr-2017	7	98
	3	4-Apr-2017	7	98
Lot 05	0	8-Dec-2016	6	96
	1	4-Apr-2017	7	96
	2	16-May-2017	7	95
Lot 07	0	3-May-2017	6	96
	1	1-Jun-2017	6	96
	2	10-Aug-2017	6	96

ASIC Lot	Aging Cycles	Collection Date	Reference Replicates	No. Samples
Lot 08	0	16-May-2017	7	116
	1	20-Jun-2017	7	116
	2	18-Sep-2017	7	113
Lot 09-12	0	12-Oct-2017	6	80
	1	4-Dec-2017	6	80
	2	5-Mar-2018	6	80
Lot 10-11	0	10-Oct-2017	6	84
	1	4-Dec-2017	6	79
	2	25-Jan-2018	6	79
Lot 14	0	14-Jun-2018	6	96
	1	16-Jul-2018	6	94
	2	24-Aug-2018	6	93
Lot 15	0	12-Jun-2020	6	94

2.2. Reference Variability

During each sample collection event, PSA spectra of the reference sample (Sample 166) were collected after every 20 spectra, resulting in 6–9 replicate reference spectra for each collection. These replicate spectra help define the variability of one sample—both between PSA collection events and within one collection.

The variability of the reference PSA spectra was visualized using t-SNE (t-distributed stochastic neighbor embedding), a method of visualizing high-dimensional data on a two-dimensional plane while preserving relative distances between points.[4] The visualizations presented here are based on Euclidean distance metrics. The t-SNE results for the raw (unnormalized) data are shown in the left panel of Figure 3, with the data colored by collection event. The spread of the data show that the PSA data vary with collection date, and the variability between dates can change dramatically with collection date.

The same data after normalization to the reference spectrum (one of the replicate spectra collected during the event) are shown in the right panel of Figure 3. The spread in the t-SNE map now appears very different; replicate spectra from each collection event (typically ~5 spectra) are clustered tightly together, and the individual collection events are fairly evenly distributed across the t-SNE map. This demonstrates that the PSA collection is stable over a collection event, and instrument variability primarily occurs between different collection events. While reference normalization appears to greatly reduce the variability across collection events, the separation of collections in the map suggests it is not a complete solution.

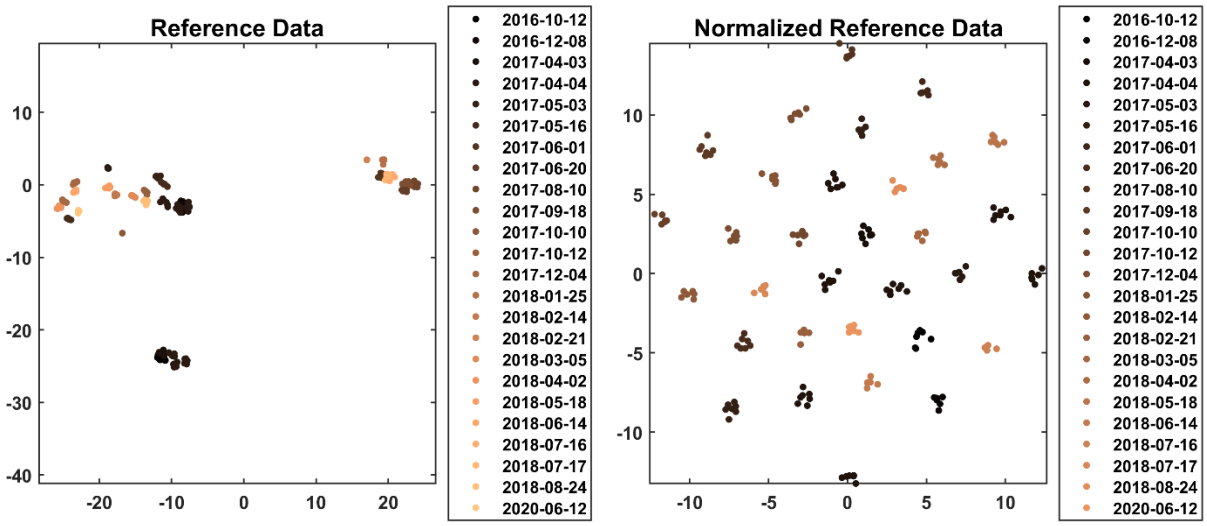


Figure 3. t-SNE results for replicate reference measurements (Left) Unnormalized (raw) PSA spectra (Right) Reference-normalized PSA spectra

2.3. Control Sample Variability

Similar analysis can be performed on the PSA data collected from ten control samples at each collection event. These samples are nominally unchanged between collection events, so are useful to explore how the PSA data vary in time. If normalization to the reference spectrum is effective in removing instrument variability, the control samples should show no changes related to sample date after normalization.

The t-SNE map of the control data before and after normalization are shown in Figure 4 and Figure 5. This time, the data were first decomposed by PCA (five components) prior to visualization using t-SNE. Prior to normalization, the t-SNE map shows strong clustering by collection date. After normalization, the data appear to cluster primarily by the control sample number. The normalization shows that a few of the control samples appear to be distinct from the others, particularly control sample 182 (Figure 5, right panel).

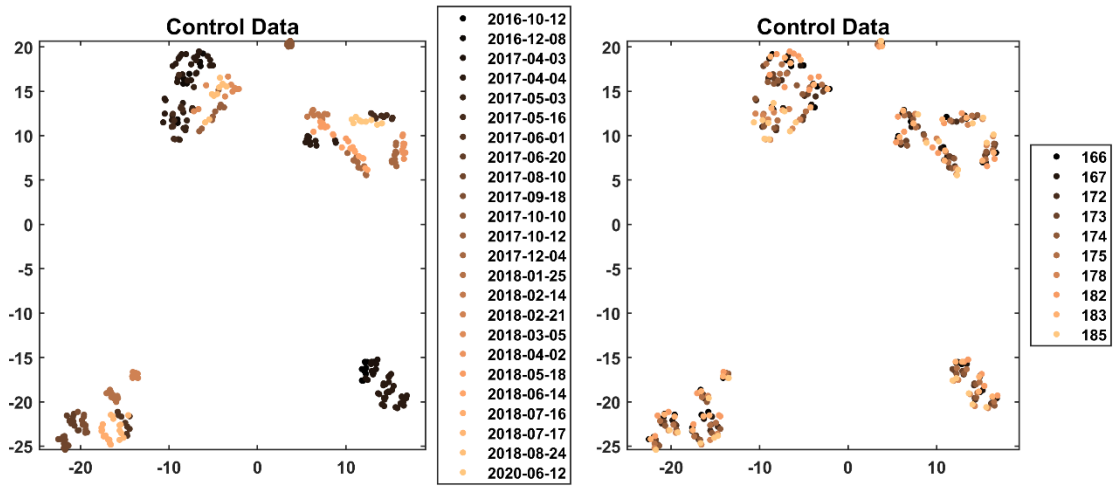


Figure 4. t-SNE results for control measurements (Left) Colored by sample date (Right) Colored by control sample

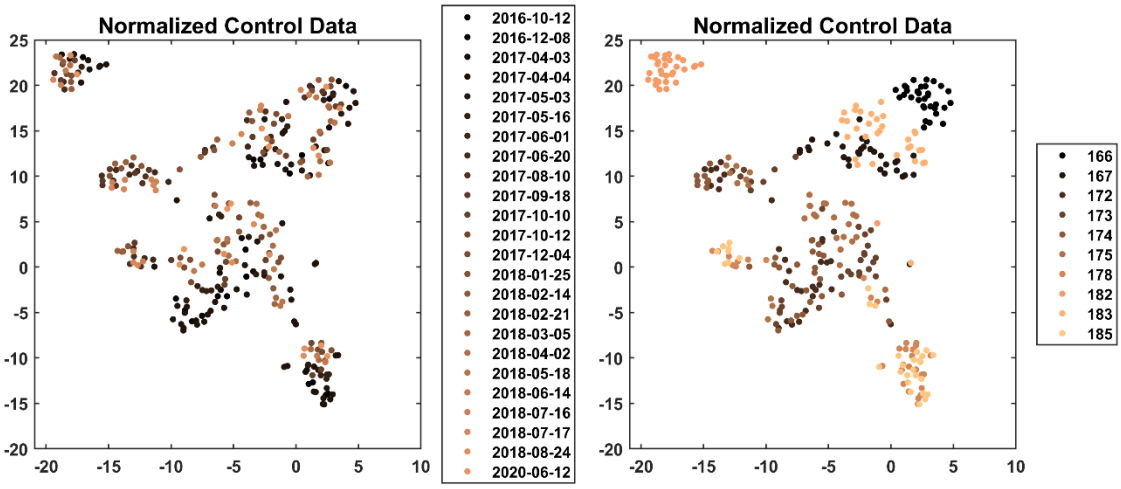


Figure 5. t-SNE results for control measurements after normalization (Left) Colored by sample date (Right) Colored by control sample

The variability of the control samples can be further investigated using PCA. All control data were modeled together using five PCs, selected by looking at the variance explained by each PC (Figure 6). Two model metrics were used to assess the variability of the control data over all data collection times: Q-residuals and Hotelling's T^2. These two metrics are commonly used in process control models to assess model performance over time.[3, 5-7]

Q-residuals reflect the variability in a data set that is not captured in a PC model, and is calculated as

$$Q_i = \sum_{j=1}^{J} e_{ij}^2$$

where e_{ij} is the residual of the ith sample after modeling with a PC model built with J components.[3] The Q-residuals can be thought of as a measure of how far the data are from the model.

The Hotelling's T^2 value reflects the distance between a given sample *i* and the class centroid, and is calculated as

$$T_i^2 = \hat{t}_i(\frac{T_g'T_g}{N-1})^{-1}\hat{t}_i'$$

where *N* is the number of samples and *T* is a *NxJ* matrix of centered sample scores from the PC model.[5] Hotelling's T^2 can be thought of as a measure of how far the data are from other data within the model space. Combined, these two metrics give an idea of the in-model (Hotelling's T^2) and out-of-model (Q-Residual) variability of each sample.

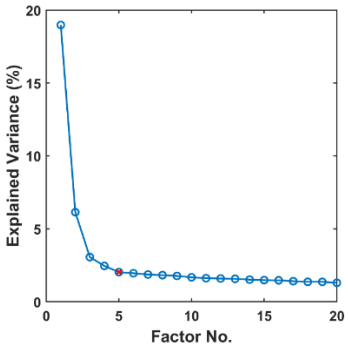


Figure 6. Explained variance in PC model of control samples

The metrics for the control samples are show in Figure 7 and Figure 8. Both figures show the reduced values, meaning the value normalized to the 95% confidence limit for the metrics. Values above one in each figure correspond to outlier samples. Notably, the samples do not show a clear trend with collection date, pointing to the stability of the control samples over time.

The Q-residuals of all samples are within the 95% confidence limit. This makes intuitive sense as all samples were used to build the model, so it is expected to be appropriate to model the data. Of more interest are the Hotelling's T^2 values, which show in-model variability. All points above the limit come from a few dates: 2018-05-18, 2017-04-04, and 2017-08-10. Because those dates show consistently higher variability across control samples, it appears that instrument changes are driving the sample variability. While most of the spread appears related to sample date, the mean Hotelling's T^2 value shows differences across control samples, showing that individual samples behave differently within the model space. These differences are consistent with those seen in the t-SNE analysis in Figure 5.

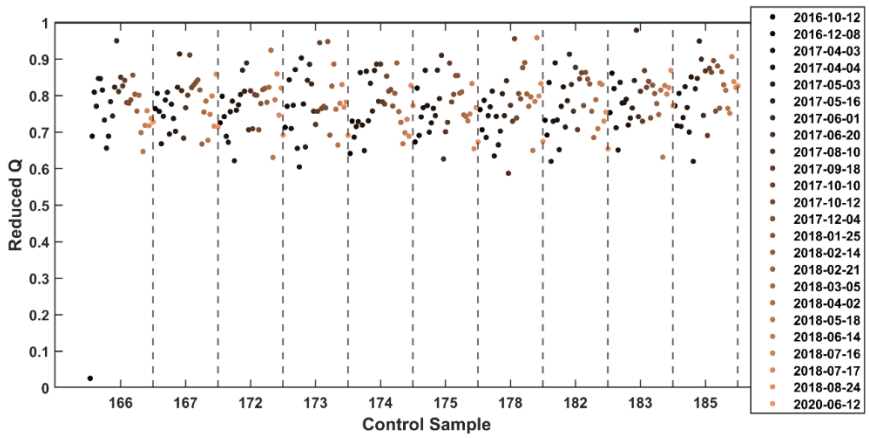


Figure 7. Control sample Q-residuals, normalized to 95% confidence limit

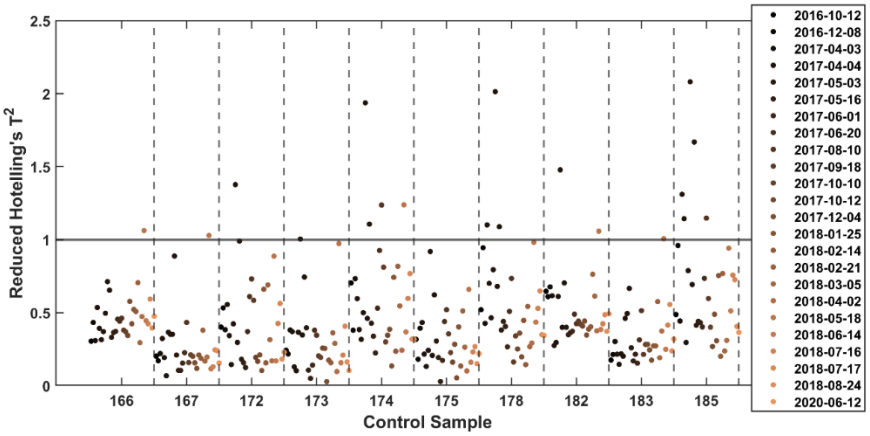


Figure 8. Control sample Hotelling's T^2, normalized to 95% confidence limit

Understanding potential influence of individual samples or instrument changes on model performance is critical to assess model performance as a screening tool that will likely be used to assess data collected over a long period of time. One way to estimate this using relatively limited data is to perform cross-validation, in which the overall model performance is compared to that achieved by using only a subset of the data. Here, the models are assessed first by leaving out individual control samples, and then by leaving out individual collection dates.

The results of leaving out individual sample dates are shown in Figure 9. The open black circles are the values obtained using the full data model. The colored circles are the results of each date when that date is removed from the model. For example, the blue Xs at 2018-05-18 are the reduced Q-residuals and reduced Hotelling's T^2 values from a model built on all data *except* the 2018-05-18 data. In that case, the Hotelling's T^2 values are much lower than they are with the full model, and the Q-residuals are higher. This shows that these data are now further from the modeled distribution (contain variability not adequately modeled) when they are not included in the model. This is also true of the 2017-04-04 and 2017-08-10 data, all dates that behave dissimilarly from the others in the original models.

The change in model performance after these samples were removed shows that models of the PSA data will be susceptible to variability with instrument changes. To reduce this influence, it is recommended that control samples be collected with each PSA collection (as they were here) and

that these data be compared to a representative distribution before additional data are collected to ensure consistent PSA data.

In cases where this is not possible, a correction method for instrument variability was assessed. This correction was performed by subtracting the mean PSA spectrum of the control samples at each date from every PSA spectrum collected on that date. The results are shown in Figure 10. This time, the results of each model were more consistent with the results of the overall model, showing increased robustness to instrument variability. The application of this correction will be discussed in the application examples below. The utility of such an approach relies on the assumption that the primary form of instrument drift will be additive, so performing a corrective shift will be adequate. This will not be true if the instrument changes include a multiplicative component; in that case, further corrections might be required. It also relies on the collection of data that are expected to remain unchanged across collection events, such as the control samples included in this study.

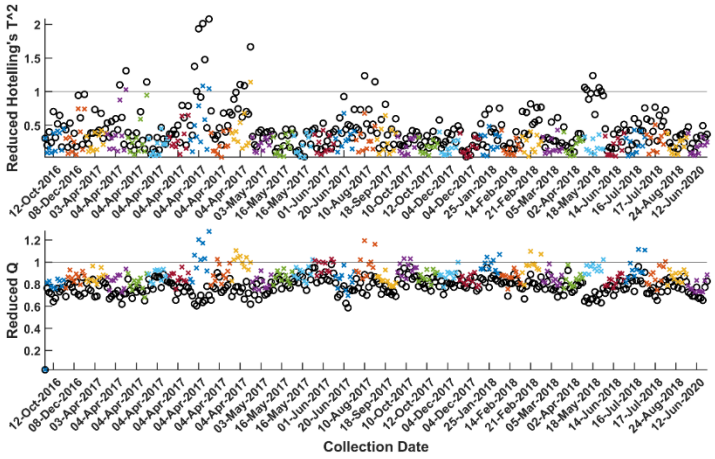


Figure 9. Model performance when leaving out individual dates

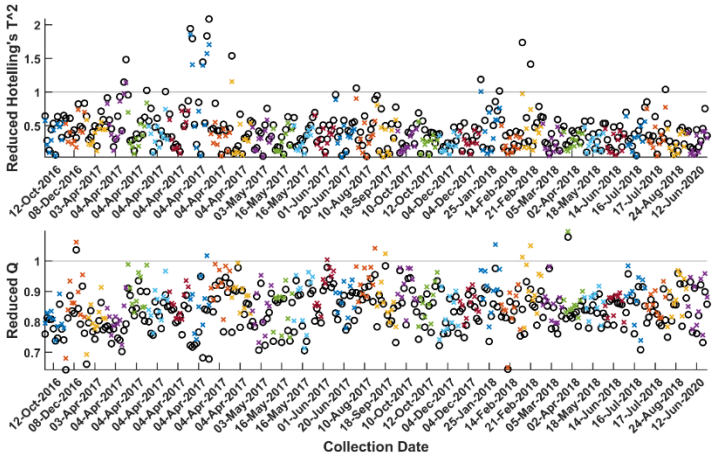


Figure 10. Model performance when leaving out individual dates after centering by the mean of control samples on each date

Similar analysis was performed by leaving out individual control samples across all collection events (Figure 11). Unlike analysis by date, removing individual samples from the models did not appear to significantly affect the model performance on those withheld samples.

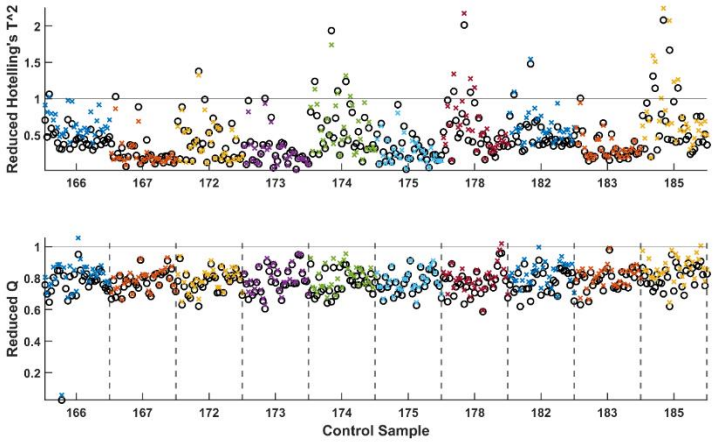


Figure 11. Model performance when leaving out individual samples

To understand the number of samples that might be needed to adequately define the sample distribution, models were developed using 1–9 control samples, employing different combinations of the original 10 control samples. Importantly, each sample included the 32 spectra from across the collection events, as collection dates appear to be the greater driver in modeling.

Results of these models are shown in the boxplots in Figure 12–Figure 13. The data shown in the right plot of Figure 12 are identical to those in the bottom panel of Figure 11. In each plot, the Q-Residuals included in each box are for all PSA spectra of that sample in all models where that sample was withheld. In the left plot of Figure 12, all 10 models were built using only one control sample, and so demonstrate the worst case for modeling additional data. The reduced Q-Residuals across all withheld data are consistently above one, showing that the model is unable to recognize similar samples. The data shown in the right plot of Figure 12 are the best case for these data, using nine of the ten control samples. Here, withheld sample spectra consistently show reduced Q-Residuals <1, and so are correctly identified as in-group. All combinations between these two extremes were evaluated (not shown), and the first number of samples that show consistently acceptable performance are six-sample models (Figure 13), suggesting that at least six independent units evaluated at multiple time points are needed to model the distribution of a similar data set (preferably with multiple spectra collected at each time).

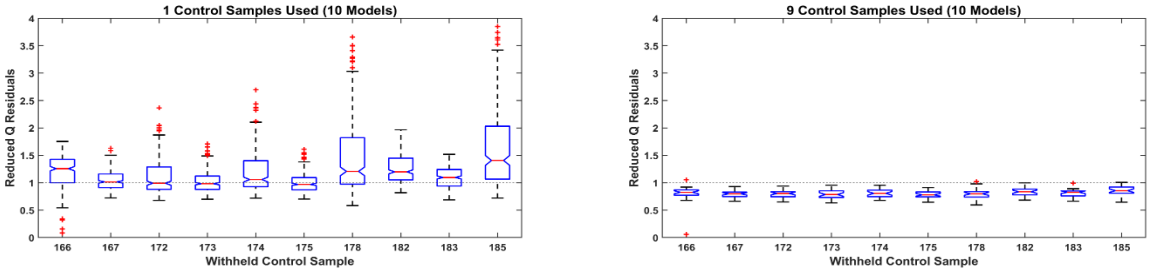


Figure 12. Reduced Q-Residuals (Left) Models built with one control sample (Right) Models built with 9 control samples

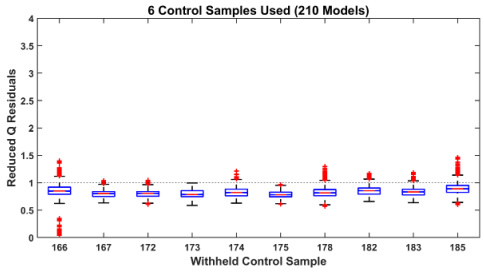


Figure 13. Reduced Q-Residuals of models built with 6 samples

2.4. Recommendations to control variability

Implementation of PSA as a larger scale screening tool requires good control of data variability over the course of continued collection. The data presented here demonstrate that spectra collected from a single unit within one collection event are consistent (Figure 3), meaning that the PSA test set-up does not exhibit significant short-term drift, and changes in sample placement between replicate measurements are not significant. Changes in the PSA experimental set-up between collection events are a much stronger source of variability. While this effect is partially compensated for by normalization to a reference sample from that collection, the variability persists through modeling and can be expected to affect future model performance (Figure 9). Further, individual sample units from a given device lot can also exhibit variability that affects model performance on new, unseen units (Figure 5).

From this analysis, a few recommendations can be made for future use of PSA to enable reliable modeling of the data:

1. Baseline models should include data collected across multiple collection events.

   It is best practice to represent any source of variability in the data expected in future data. In this case, variability of the PSA data appears to be dominated by variability between collection events, or variability in the instrument set-up. Therefore, models that encompass several collection events are expected to be more robust to variability of future collection events.

2. Collect control samples with each collection event and compare these control samples to expected performance prior to collecting any further data.

   Control samples (unaged, known good parts) are typically collected alongside test samples of interest; this practice provides a good understanding of any potential changes in the instrumental set-up that might influence model performance. This practice can also be improved by maintaining a control model that can be used to perform a quality check prior to collection of any sample data. As long as the control samples behave as expected in this quality model, it can be reasonably expected that any sample data collected would be representative of true changes in sample units, rather than changes in the experimental set-up.

3. Multiple independent samples need to be used to develop the baseline model.

   The data presented in Figure 12 and Figure 13 show that the data collected from only one unique sample is not representative of other sample units. At least six units (each with spectra across 32 collection events) were needed to build a PCA model that was reliably representative of the other four units with data across those same collection events. The number of units needed to develop an initial model will be dependent on the variability of the units under examination, but these data point to a minimum of six units needed.

3. AGING OF ZENER DIODES

3.1. Experimental

Three groups of Zener diodes were subjected to accelerated aging conditions (recorded in Table 2). After each aging cycle, select units were chosen for destructive physical analysis. PSA data were collected for each data set at 0.08 V (forward-biased), 1V (forward-biased), and 3V (reversed-biased). Conventional current-voltage (IV) data were also collected for each device. Each group of diodes included equal samples from two different date codes. Along with the data of aged devices, PSA spectra and conventional data were collected from ten control (unaged) samples at each collection event. These controls are Zener diodes from the same date codes and were not subjected to any aging conditions.

Table 2. Aging information for Zener diodes

Group	Aging Conditions	Aging Cycle (weeks)	Samples Measured
1	110°C, 85 %RH	0	100
		1	98
		2	98
		3	98
		4	96
2	140°C	0	100
		1	95
		2	94
		3	95
		4	95
3	150°C, 65 %RH	0	90
		1	90
		2	89

3.2. Partial Least Squares Regression Overview

Partial least squares (PLS) regression (PLSR) is a multivariate analysis technique that decomposes a data matrix into factors that maximize the explained covariance between the data set and a known metric. In this case, the data matrix is either PSA or conventional data, and the known metric is the sample age, in weeks or cycles. This technique is particularly powerful when the data matrix has a large number of correlated variables. The resulting factors, or latent variables (LV), of the PLS model identify the correlated data features most closely related to the sample age.

The details of PLSR are thoroughly covered in the references. [8-11] Briefly, score and loading matrices are estimated for both X (variable matrix) and Y (response matrix) data. The loadings, analogous to spectra, describe the relative contribution of each variable within the LV, while the scores describe the relative contribution of each LV to the overall data signature. During the

estimation of these scores and loading, a matrix of weights is also obtained that aids in defining the relationship between the X and Y data. Ultimately, a regression vector *b* is obtained such that

$$Y = Xb + E.$$

Values of *y* for new samples can then be estimated as

$$\hat{y}_i = x_i b.$$

The regression vector is a weighted, linear combination of the loading vectors that weights the contribution of each variable to estimate $\hat{y}$. The values of $\hat{y}$ will be the age estimate in Section 3.4.1, or they can be compared to a metric to make a group/age determination, as in 3.3.

3.3. Multivariate Analysis of Conventional Data

Using multivariate modeling methods successfully developed and demonstrated in a previous LDRD [12] and detailed in the open literature articles,[13-20] conventional data collected on the COTS components throughout the course of the study were investigated to determine if automated algorithms could be developed to detect part aging. The PLSR modeling in this section was performed using the commercially available software, The Unscrambler®X by Camo Analytics (now owned by Aspen Tech).

Voltage data collected from the Zener diodes as a function of current were assembled into a 1×1089 matrix for each part tested to create a "signature" of the unit being tested (see Figure 14, Figure 15, and Figure 16 for plots of the conventional data). Each stress condition was analyzed by a separate algorithm (Table 2).

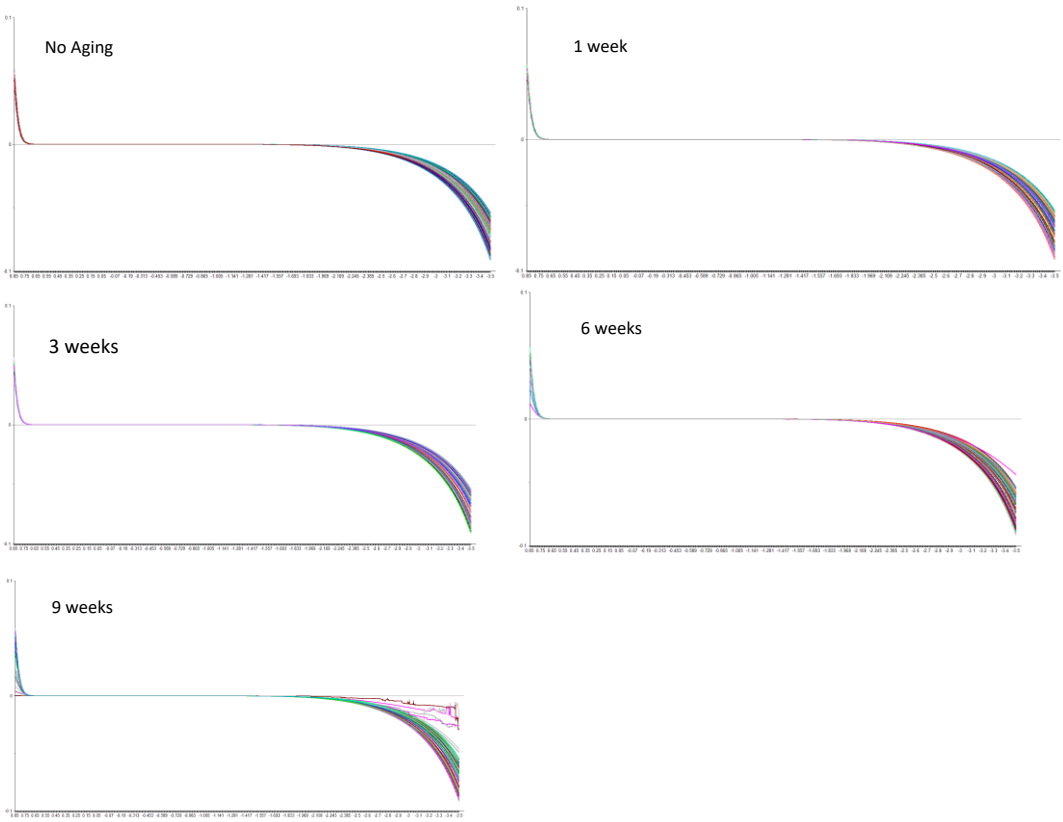


Figure 14. Plots of conventional data analyzed for Group 1

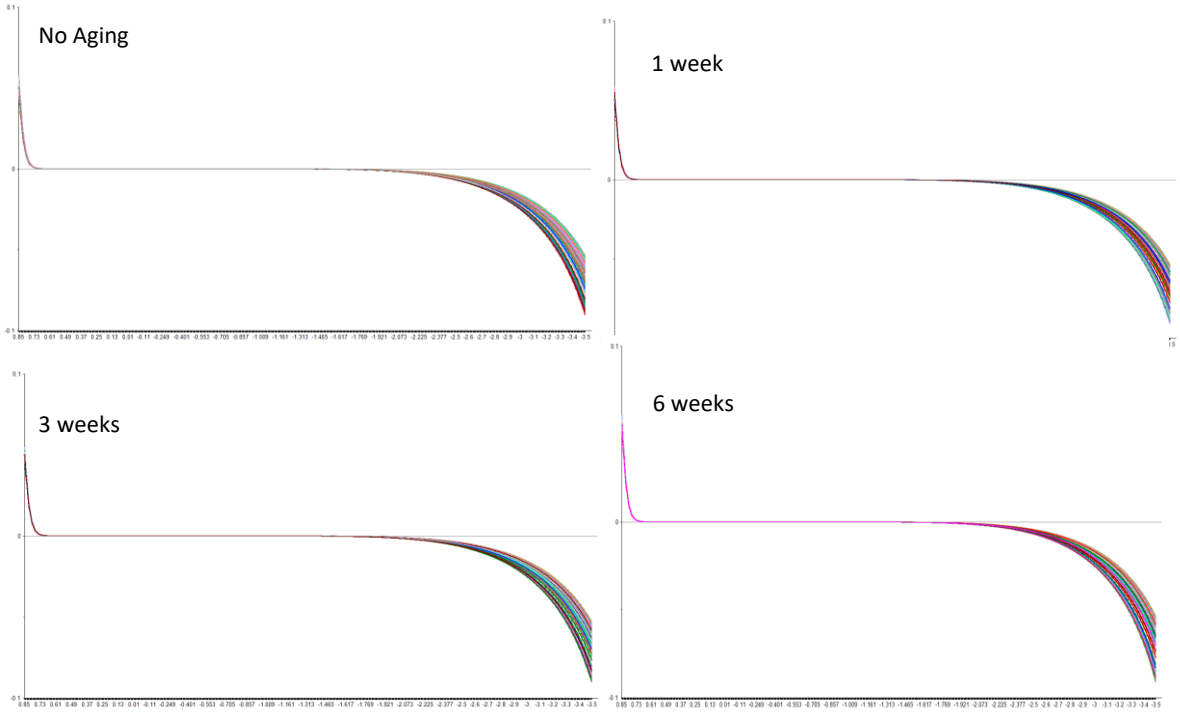


Figure 15. Plots of conventional data analyzed for Group 2

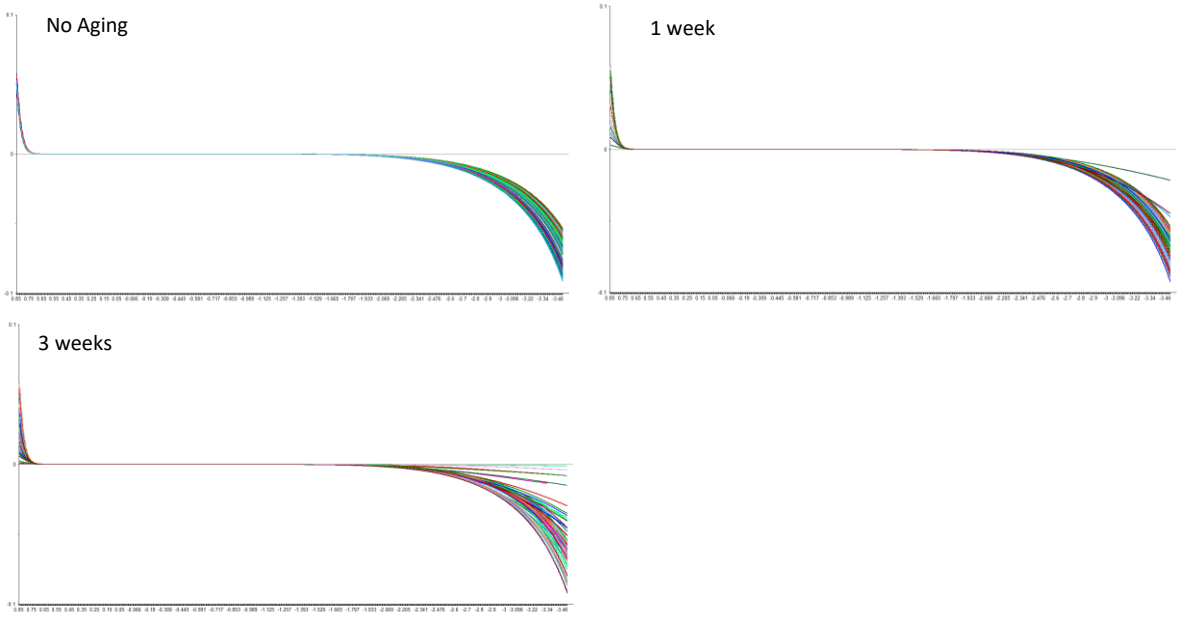


Figure 16. Plots of conventional data analyzed for Group 3

The primary assumption in creating each model is that any variation observed in the IV curves (1×M data, where M is the number of variables) can be attributed to information in the data set. Each model is developed using PLS, which transforms the data to new orthogonal axes centered on the mean data set of the data groups with axes oriented along the directions of maximum variation

of the groups (latent variable space). This transformation is applied to both the observations and responses in the data sets simultaneously, resulting in separation between the groups.

To develop algorithms to detect aging, the following process was used:

1. Data collected for each unit tested were arranged in a 1×M matrix to create a signature of the measurement for modeling as discussed above.
2. All measurement signatures were grouped by type and arranged into a matrix for regression analysis.
3. LVs were calculated for the analysis matrix by modeling both the X and Y matrices (variables and responses) simultaneously using known data via PLSR.
4. The optimal number of LVs was selected to best separate the groups in the modeling. These LVs were used to estimate a regression vector *b* as described in Section 3.2. This vector has a characteristic linear equation to relate variables to the response

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \cdots + b_M x_M.$$

5. The value of $\hat{y}$ can then be used to indicate how well-matched new input data are to one of the response groups in the modeling.
6. Multiple models were combined to create a programmed flow to differentiate new data based on $\hat{y}$ values determined by each model.

The algorithms developed following this methodology are outlined in Table 3. Figure 17, Figure 18, and Figure 19 show the results for Group 1, Group 2, and Group 3 algorithms respectively when tested on new input data not included in the modeling.

Table 3. Models of Conventional Data

Group	Calibration Data	Validation Data	Results on Validation Data		Notes
Group 1	59	597	6 Weeks	95% Correct (82/86 parts)	Note: Conventional data from the no-aging units collected after 10/2021 could not be differentiated from the aged units in the Group 1 algorithm that were created and tested using all data. This indicates that it is different from data collected prior to 10/2021.
			9 Weeks	91% Correct ID (82/90 parts)	
			0–3 Weeks	Poor differentiation 63% of no-aging units correctly identified (122/193) 92% of 3 week units correctly identified (166/179)	
Group 2	66	517	6 Weeks	96.6% Correct ID (85/88 parts)	

Group	Calibration Data	Validation Data	Results on Validation Data		Notes
			1 Week	100% Correct ID (87/87 parts)	22.6% of no-aging parts were misidentified as 3 weeks (56/248 parts)
			3 Week	91.1% Correct ID (72/79 parts)	
Group 3	28	415	1 Week	85.1% Correct ID (74/87 parts)	
			3 Week	89.7% Correct ID (78/87 parts)	
			0 Week	85.9% Correct ID (207/241 parts)	

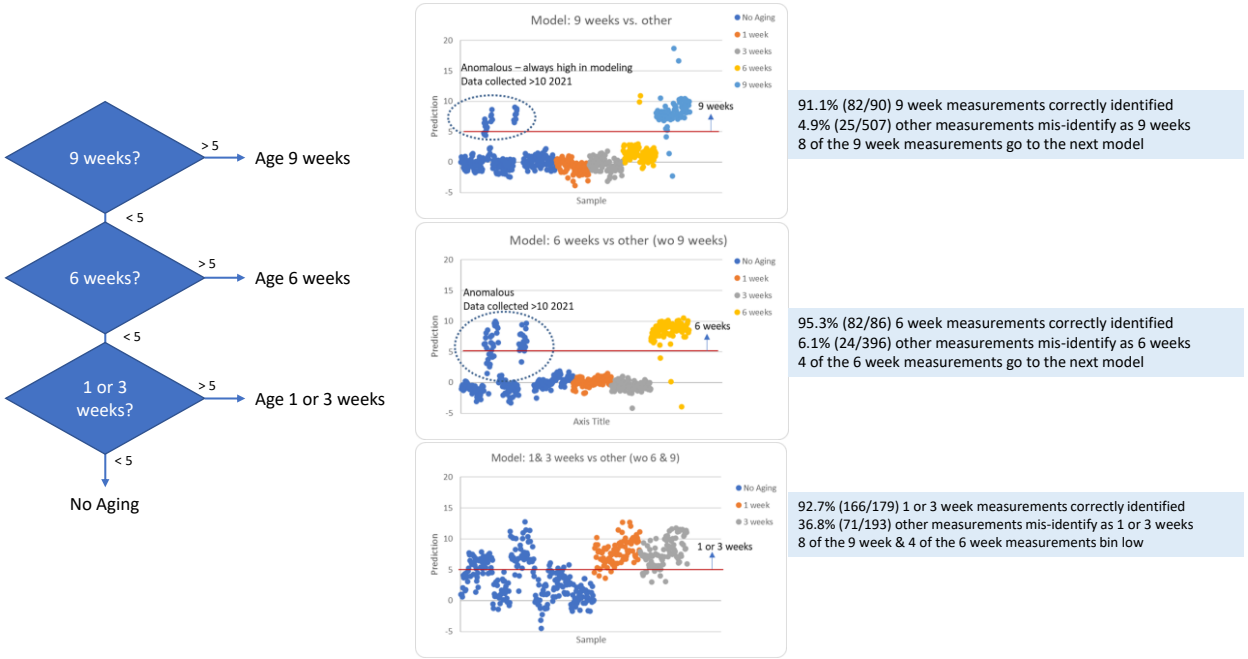


Figure 17. Results for the Group 1 Algorithm to identify aged parts when tested on new input data not included in the modeling. Note: No-aging part data collected post 10/2021 presents as aged.

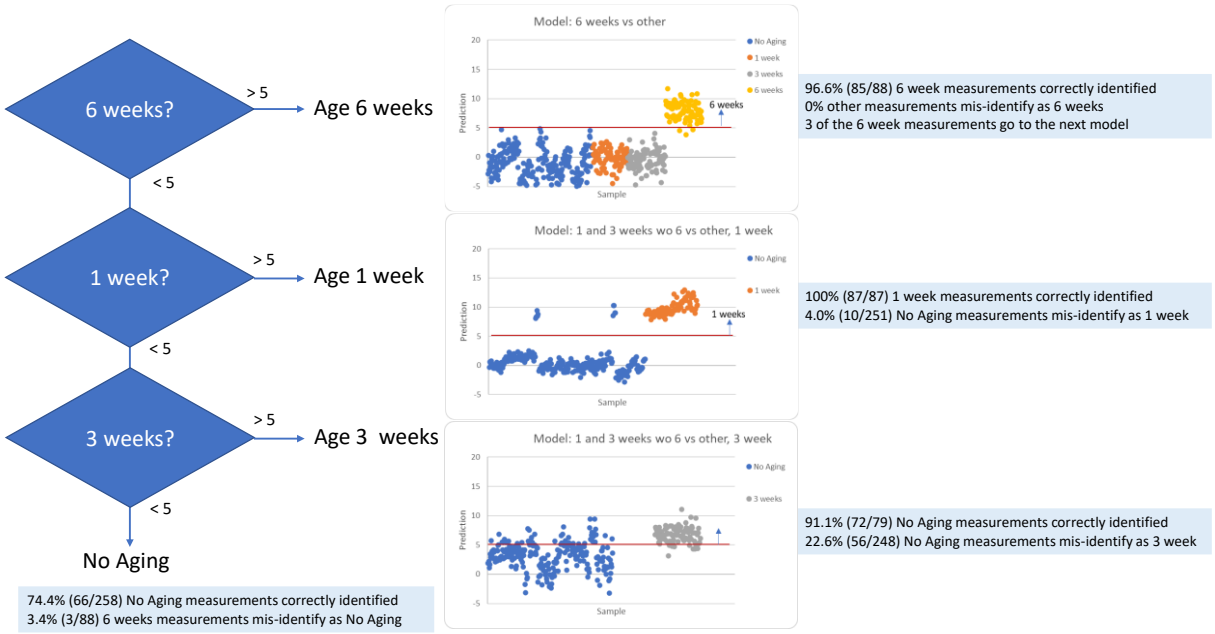


Figure 18. Results for the Group 2 Algorithm to identify aged parts when tested on new input data not included in the modeling.

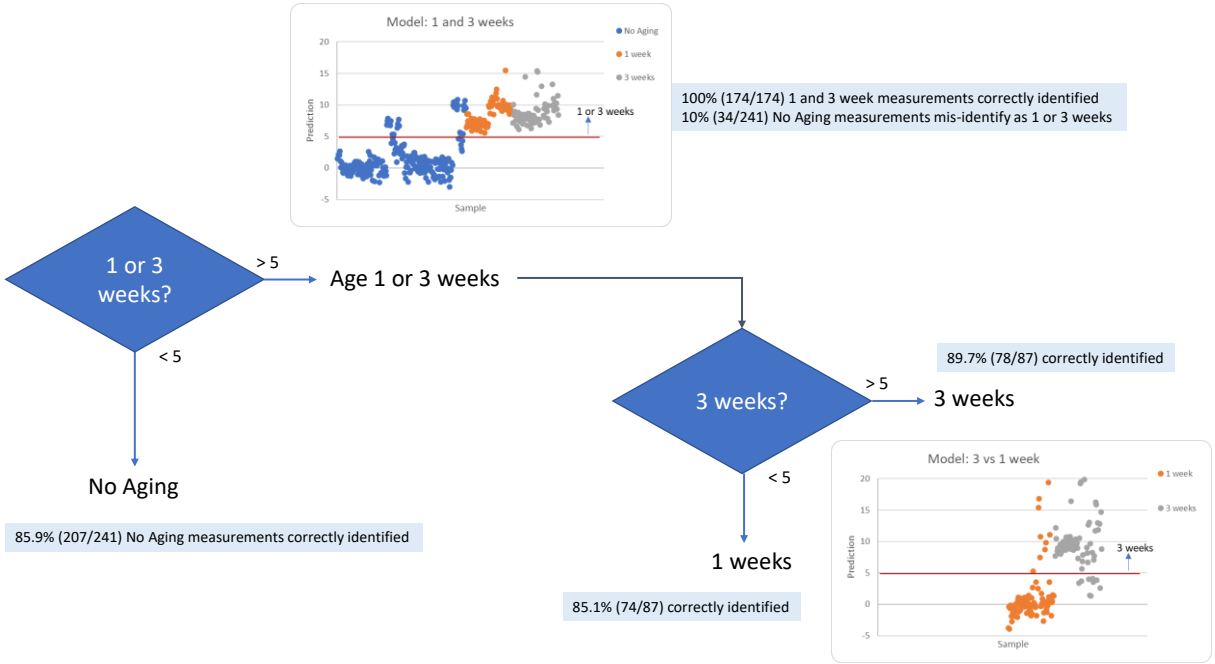


Figure 19. Results for the Group 3 Algorithm to identify aged parts when tested on new input data not included in the modeling.

The modeling of Group 1 data shows that the data collected after 10/21 differ substantially from data collected before that date. This is likely related to instrumentation or other test-set up changes, causing difficulty in creating predictive models. To investigate the effects of these changes, a second Group 1 model was created using only data collected after 10/2021. This model demonstrated that

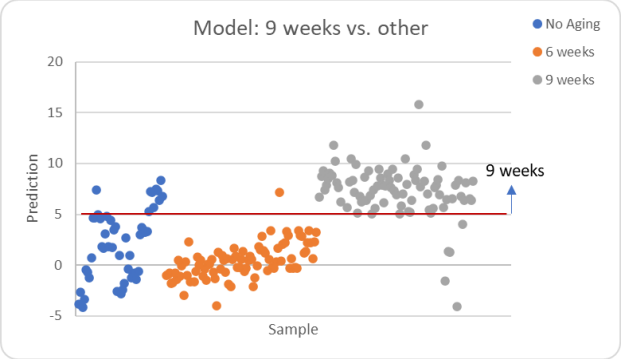
data collected after nine and six weeks of aging were differentiable from no-aging data (Table 4, Figure 20).

Table 4. Second Model of Group 1 Conventional Data

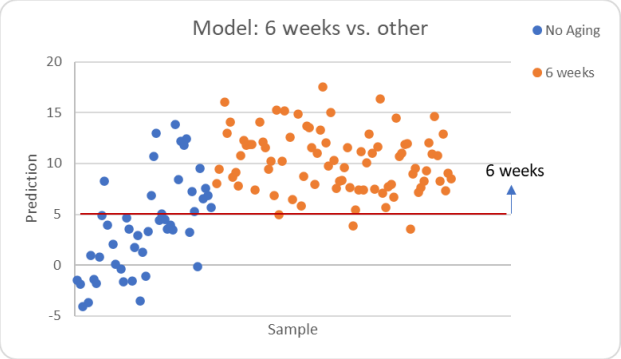
Group	Calibration Data	Validation Data	Results on Validation Data		Notes
Group 1: Model of data after 10/2021	25	228	9 Weeks	89.9% Correct ID (80/90 parts)	4.8% (11/228) 0 and 6 week samples identified as 9 weeks 7.5% (17/228) 0 and 9 week samples were identified as 6 weeks
			6 Weeks	90.9% Correct ID (80/88 parts)	
			0–3 Weeks	Not differentiable	

The algorithm's ability to differentiate nine- and six-week data from no-aging data validates that conventional data and multivariate analysis can recognize aging parts.

Conventional data from the no-aging units collected after 10/2021 could not be differentiated from the aged units in the Group 1 algorithm that were created and tested using all data. This indicates that it is different from data collected prior to 10/2021.



89.9% correct ID (80/90 parts) as 9 weeks
4.8% (11/228) mis-ID as 9 weeks



90.9% correct ID (80/88 parts) as 6 weeks
7.5% (17/228) mis-ID as 6 weeks

Figure 20. Results for the models created using Group 1 data collected 10/21 or later to identify aged parts when tested on new input data not included in the modeling.

In summary, conventional data combined with multivariate analysis can be used to create algorithms to screen COTS parts for aging with ~90% or better correct ID, depending on the stress condition for 9-, 6-, 3-, and 1-week-aged parts. In addition, from the behavior of the Group 1 data, something changed in the conventional data set-up or units for parts in this group tested 10/2021 and later.

3.4. Models to Estimate Age from PSA Spectra

3.4.1. Method 1: Partial Least Squares Regression Model of All Aging Data

Partial least squares regression (PLSR) was used to model the PSA data from Group 1 Zener diodes. Prior to building the PLSR model, outliers were identified and removed from the data set, based on the 98% confidence limits of the Q-residuals and Hotelling's T^2 values of an initial PLS model built with all of the data. 58 spectra (~10%) were identified as outliers. Of the remaining data, 30% were withheld as validation samples (130 samples) and 70% were used for calibration (303 samples). In addition, the control samples collected at each time were used as validation. The rank of the model was identified using venetian blind cross-validation with five sample splits, resulting in ~20% of calibration withheld in each cross-validation iteration. The PSA data were mean-centered and the aging cycle data were auto-scaled prior to calibration.

An initial visualization of the Group 1 Zener diode data is shown in Figure 21. The plot of PC1 and PC2 shows two large clusters in the data; these clusters correspond to the two date codes of the samples. The unaged data appear to have a slight separation from the aged data. In this space, it would be expected that the control data (black diamonds) would appear near one another, and near the unaged data. Instead, the control samples appear close to the aging samples that were collected in the same collection event.

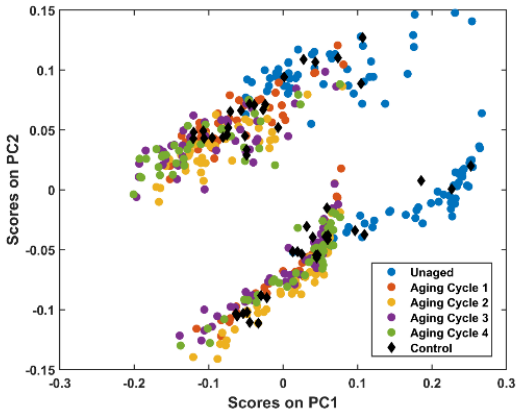


Figure 21. PCA of Zener Diode Group 1 PSA Data

This behavior is also seen in initial PLS models of the data. The model errors used to select the five LV used in the model are shown in the left panel of Figure 22. The model errors include the root-mean-squared error of calibration (RMSEC), root-mean-squared error of prediction (RMSEP), and root-mean-squared error of cross-validation (RMSECV). These errors assess model performance of the modeled data, the withheld validation data, and a set of iteratively withheld data (or cross-validation samples), respectively. The model predictions are plotted against the actual aging cycle in the right panel. The dashed line on that plot is the 1:1 line to visually show where a perfect prediction model would follow. While the calibration and validation data show good prediction of the aging cycle from the PSA data, the similar prediction of the control data clearly show that the model is reflective of sample collection date rather than aging cycle. If the model were truly reflecting aging cycle, the control samples would predict at roughly the same value across the entire model space.

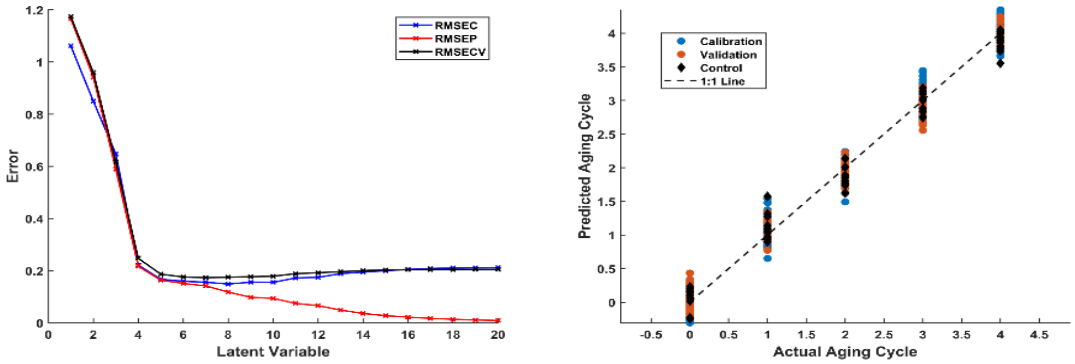


Figure 22. PLS Model (Left) Errors (Right) Prediction

To correct for the influence of collection date on the data, the mean of the control samples collected on each date was subtracted from the PSA data collected on that date. Now, the PC plot of the centered data shows the same separation of date codes, but less separation of the control samples (Figure 23).

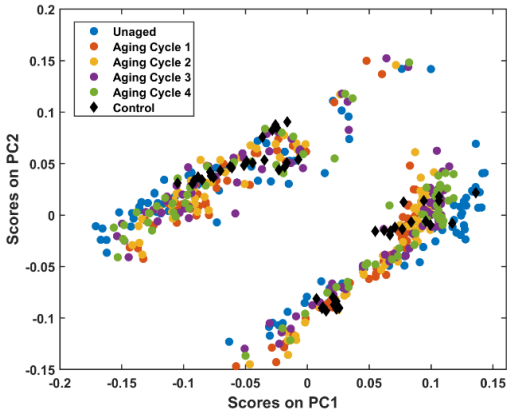


Figure 23. PCA of Centered Zener Diode Data

The PLS model of the centered data is shown in Figure 24, now modeled with six latent variables. The control samples now appear at a consistent value across the model space, showing that the model is no longer simply modeling the change in instrument set-up over time. This model shows prediction of the aging cycle from the PSA data. The data scores at each aging cycle contain a large spread in the model space, resulting in uncertainty in model predictions. This might be improved by incorporating additional data in the model.

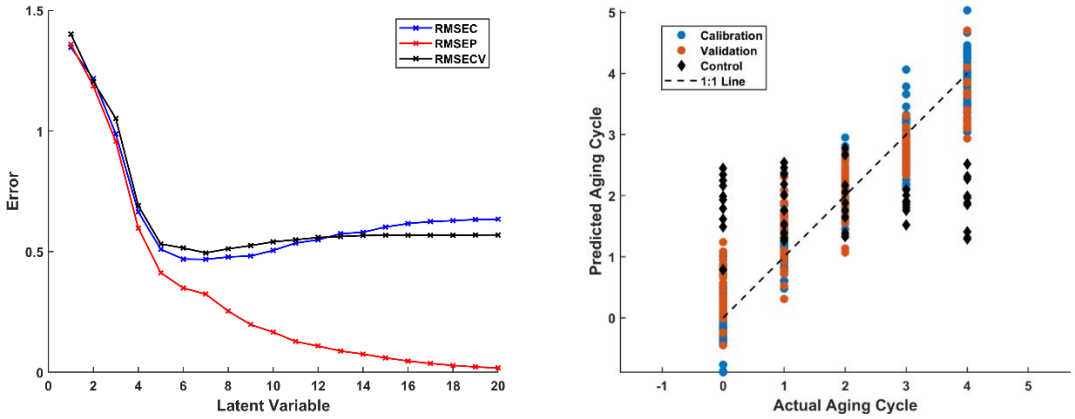


Figure 24. PLS Model of Centered Data (Left) Error (Right) Prediction

3.4.2. Method 2: Sequential Partial Least Squares Classification Algorithms

Using the methods of algorithm building described in Section 3.3, Group 3 PSA data were analyzed to determine if an algorithm could be created to predict aging cycle. The PSA data for each measurement were mean centered and scaled by the spectrum standard deviation prior to analysis (Figure 25). By applying the previously described algorithm building methods, it was possible to build an algorithm to correctly predict aging cycle by modeling on 79 (Cycle 1 and 0) or 111 (Cycle 2, 1, and 0) randomly selected measurements. The algorithm was tested on 1,476 other measurements not included in the modeling and it was found that Cycle 2 measurements could be differentiated from Cycle 1 and Cycle 0 measurements to 91.7% (289/315) accuracy with 8.3% (26/315) mis-ID; Cycle 1 measurements could be differentiated from Cycle 0 measurements with 96.5% (307/318) accuracy and 3.5% (11/318) mis-ID; and 97.5% (930/954) Cycle 0 measurements could be correctly identified with 2.5% (24/954) mis-ID. See Figure 26 for the algorithm flow and analysis results.

Table 5. Models of Group 3 PSA Data

Group	Calibration Data	Validation Data	Results on Validation Data	
Group 3	79 Cycle 2 111Cycle 1, 0	1476	Cycle 2	91.7% (289/315) Correct ID 8.3% (26/315) mis-ID
			Cycle 1	96.5% (307/318) Correct ID 3.5% (11/318) mis-ID
			Cycle 0	97.5% (930/954) Correct ID 2.5% (24/954) mis-ID

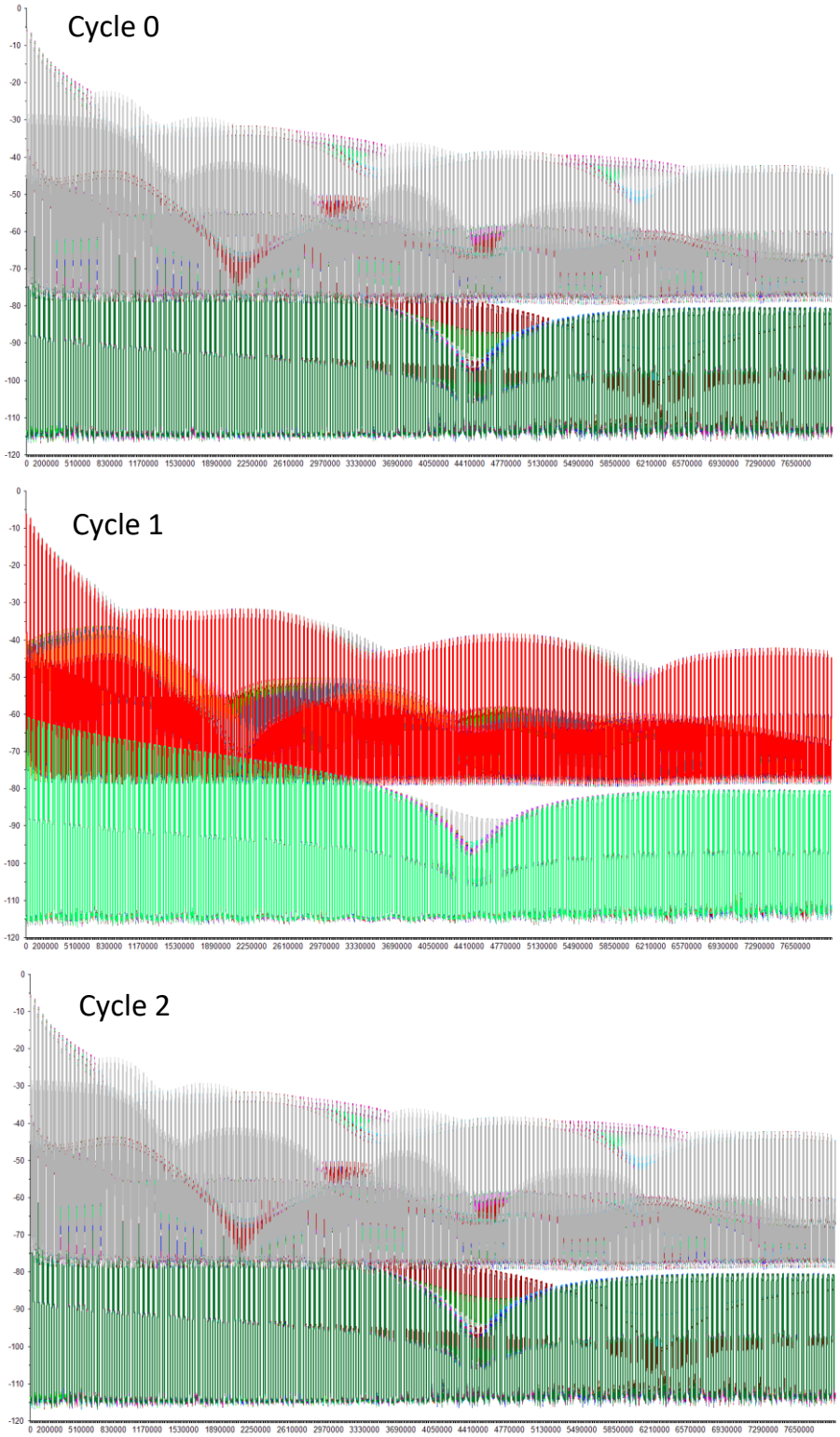


Figure 25. Plots of centered and scaled PSA data analyzed for Group 3

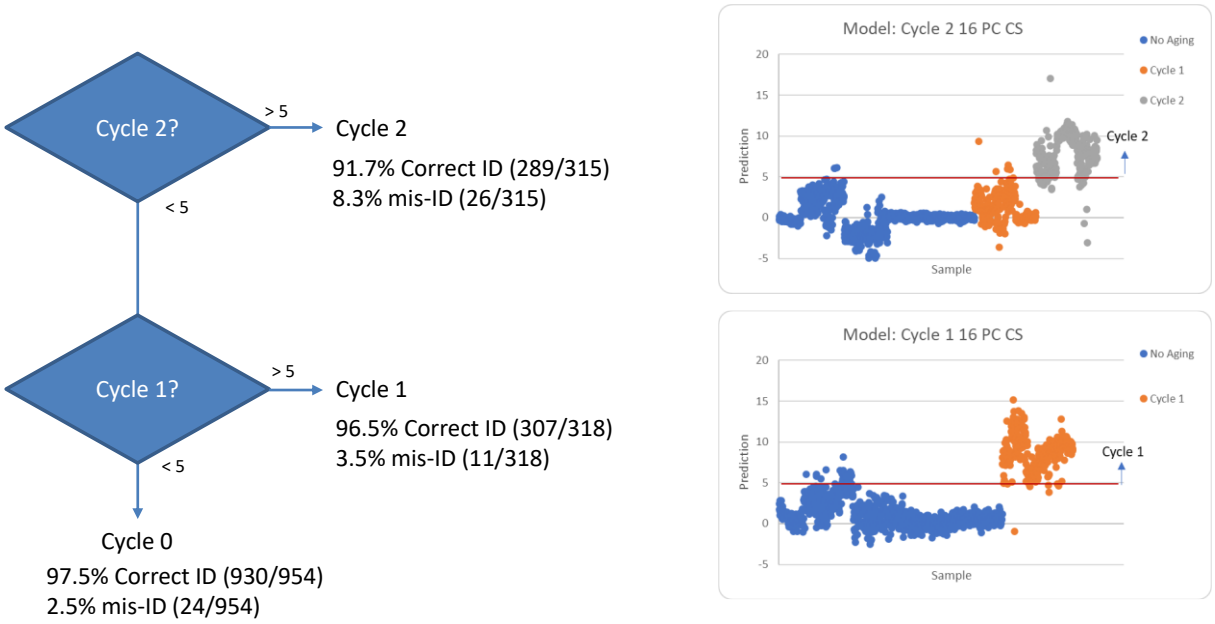


Figure 26. Results for the Group 3 PSA Algorithm to identify aged parts when tested on new input data not included in the modeling.

This page left blank

4. CLASSIFICATION MODELS

4.1. General Approach

To deploy PSA as an acceptance screening tool, data analysis must be paired with an automated decision-making algorithm to determine accepted/rejected units. Automation allows fast decision making and reduces operator bias in making determinations. One-class classification tools are useful in situations such as acceptance screening and anomaly detection, where the interest is in identifying "good" or in-class samples while the possible range of "bad" or out-of-class states are varied and largely unknown. In the face of this uncertainty, one-class classifiers depend only on correctly defining the in-group, and new samples are assessed for membership of that class.

One such approach is based on quadratic discriminant analysis (QDA), extended for application after PCA.[3, 5, 6] For PSA data, PCA is first employed to reduce the dimensionality of the highly correlated data. In this case, disjoint PCA is employed: separate PCA models are developed for each class. Specifically, data of only one class—acceptable units—are used to create the original PCA model. Extending this model to additional classes would require first creating a separate PCA model of the new class, making this method easily adapted to new data.

Once the baseline PCA model is developed, the Q-residuals and Hotelling's T^2 values are used to assess class membership (discussed in Section 2.3). These metrics can be used separately to identify anomalous units, or they can be combined into another metric as

$$G_i = \sqrt{\frac{Q_i}{Q_{lim}} + \frac{T_i^2}{T_{lim}^2}}$$

where Q_{lim} and T^2_{lim} are the confidence limits for each metric.[3, 5, 6] Normalizing each metric by the confidence limit brings them to a comparable scale and creates a baseline threshold for class membership at $\sqrt{2}$.

4.2. SNL ASIC Radiation

PCA models were developed for the EC Lot data from Table 1. Baseline PSA data were collected for all samples at 0.65, 1, and 3 V. Samples were then exposed to varying levels of radiation and subjected to subsequent aging cycles. PSA data were collected at baseline, after radiation, and after each of three aging cycles. Control samples were collected alongside the sample data at each collection event; these control samples were from another lot, so were not expected to model with the baseline data, but were still useful to understanding variability in the PSA experimental set-up. Following the approach in Section 3.4, the data at each aging time were centered by the control data at that date prior to PCA.

Independent models were first developed for each voltage, then the separate PSA spectra for each voltage were concatenated to form a joint model. The results for each model are shown in Figure 27–Figure 30. Summary performances are shown in the confusion matrices in Table 7, which report the number of true/false positives/negatives for each model. Better performing models will have higher rates of true positives/negatives, and lower rates of false positives/negatives. The location of each number is shown in Table 6, and throughout this report the false positive/negative entries are shaded in pale gray.

Table 6. Example Confusion Matrix

		Predicted Pristine	Predicted Anomalous
Example Model	Actual Pristine	**True Negative**	**False Positive**
	Actual Anomalous	**False Negative**	**True Positive**

For all models, the samples exposed to radiation are readily identified by G at a threshold of $\sqrt{2}$. The cleanest separation is seen in the 0.65V data set. During PSA analysis, the voltage level used to probe the sample targets different parts of the component. For these devices, 0.65V targets the package level (e.g., the filtering capacitors of the package), while 1V and 3V target the die level (e.g., the diodes, transistors and interconnects of the die). The larger G value obtained using the 0.65V data demonstrates that the capacitors of ASIC packages are the most susceptible to radiation damage.

In each model, the G values of irradiated samples appear to decrease with increasing aging cycle, with the maximum value occurring immediately after radiation exposure. This trend contrasts with the expected behavior of increased distance from baseline samples with aging. The value of G is based on a distance metric, suggesting that samples are returning to a normal state after several aging cycles. This is substantiated by the physics of these devices. It is known that exposing irradiated ASIC devices to temperatures after radiation can anneal some of the damage from radiation; this shows that PSA data are able to track those underlying physical phenomena in the ASIC devices.

None of the four models were able to reliably differentiate units that had undergone accelerated aging from pristine units using the G metric. However, the aged units can be visually identified using the Q-residuals alone, suggesting that combining the Q-residuals and Hotelling's T^2 value might reduce the strength of the metric in cases where the PSA spectra are more similar. This will be discussed further in the next section. These data again show the difficulty in separating variability of interest from variability related to the instrument set-up; in all bias models, the samples that were not aged or irradiated ("validation" samples, red circles with blue outline) appear separately from the baseline units collected on another date. This separation remains even after correcting by the mean of the control data, suggesting the correction is incomplete.

Table 7. Confusion Matrices for EC Lot PSA Data, Test Samples

		Predicted Pristine	Predicted Irradiated
3V Data	Actual Pristine	**77**/80	**3**/80
	Actual Irradiated	**0**/240	**240**/240

	Predicted Pristine	Predicted Aged
Actual Pristine	**18**/20	**2**/20
Actual Aged	**59**/60	**1**/60

		Predicted Pristine	Predicted Irradiated
1V Data	Actual Pristine	**72**/80	**8**/80
	Actual Irradiated	**0**/240	**240**/240

	Predicted Pristine	Predicted Aged
Actual Pristine	**20**/20	**0**/20
Actual Aged	**52**/60	**8**/60

		Predicted Pristine	Predicted Irradiated
0.65V Data	Actual Pristine	**76**/80	**4**/80
	Actual Irradiated	**0**/240	**240**/240

	Predicted Pristine	Predicted Aged
Actual Pristine	**17**/20	**3**/20
Actual Aged	**59**/60	**1**/60

		Predicted Pristine	Predicted Irradiated
0.65, 1, and 3V data	Actual Pristine	**78**/80	**2**/80
	Actual Irradiated	**0**/240	**240**/240

	Predicted Pristine	Predicted Aged
Actual Pristine	**20**/20	**0**/20
Actual Aged	**58**/60	**2**/60

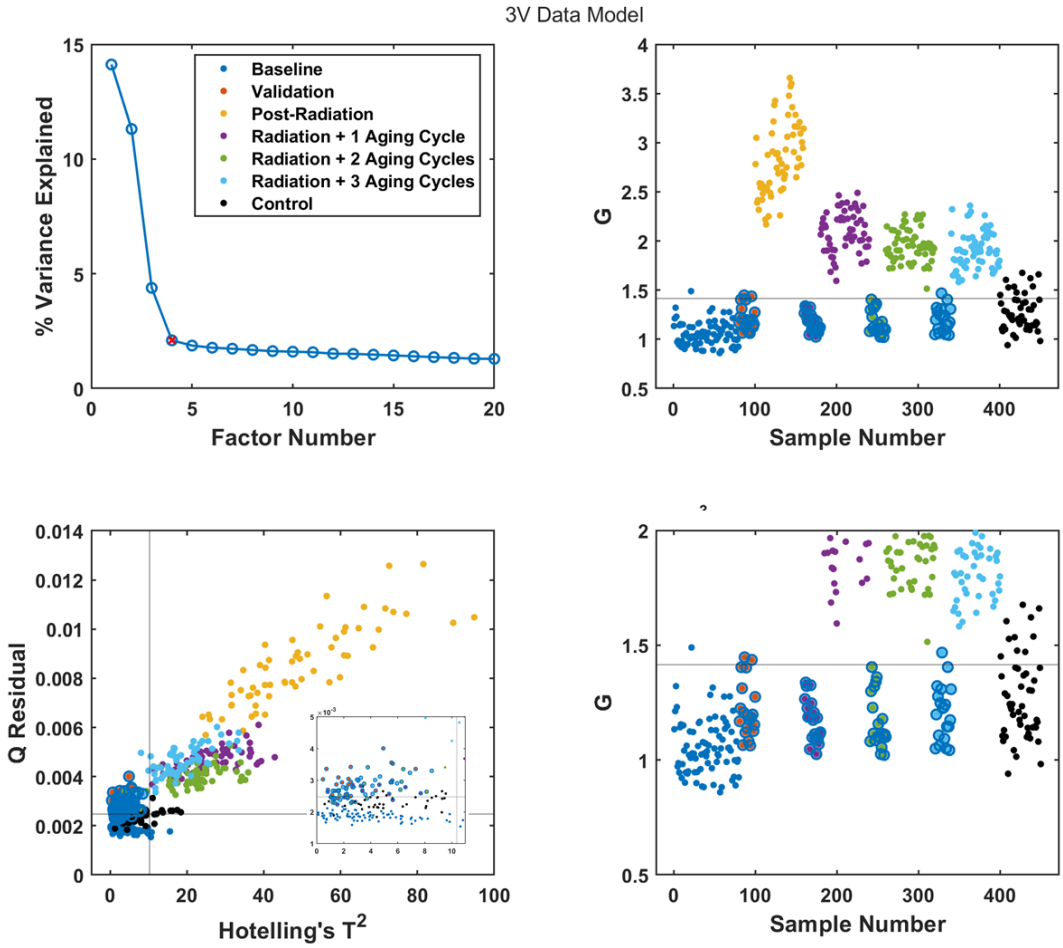


Figure 27. Classification model using 3V data (A) % Variance Explained with model rank marked (red "x") (B) G metric (C) Q-Residuals and Hotelling's T^2 (D) G metric, expanded view. Samples that have undergone aging but were *not* exposed to radiation are marked with a blue outline.

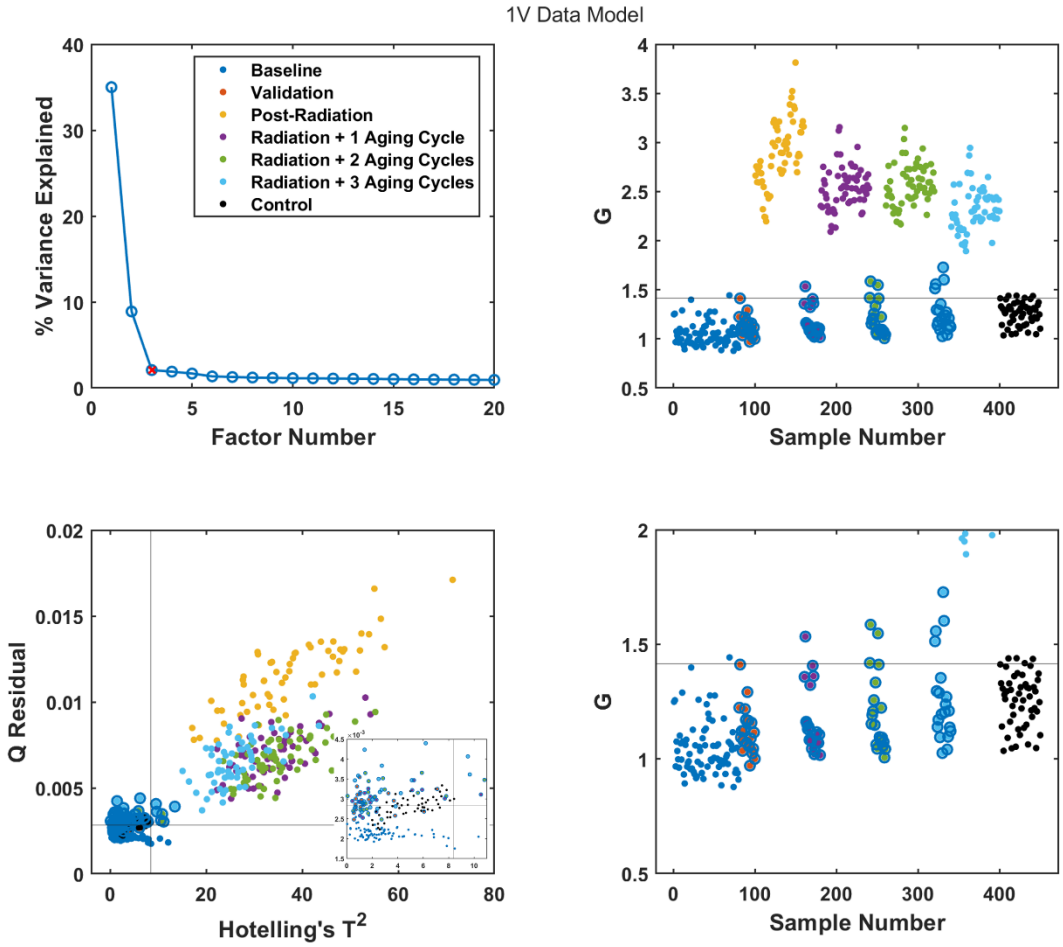


Figure 28. Classification model using 1V data (A) % Variance Explained with model rank marked (red "x") (B) G metric (C) Q-Residuals and Hotelling's T^2 (D) G metric, expanded view. Samples that have undergone aging but were *not* exposed to radiation are marked with a blue outline.

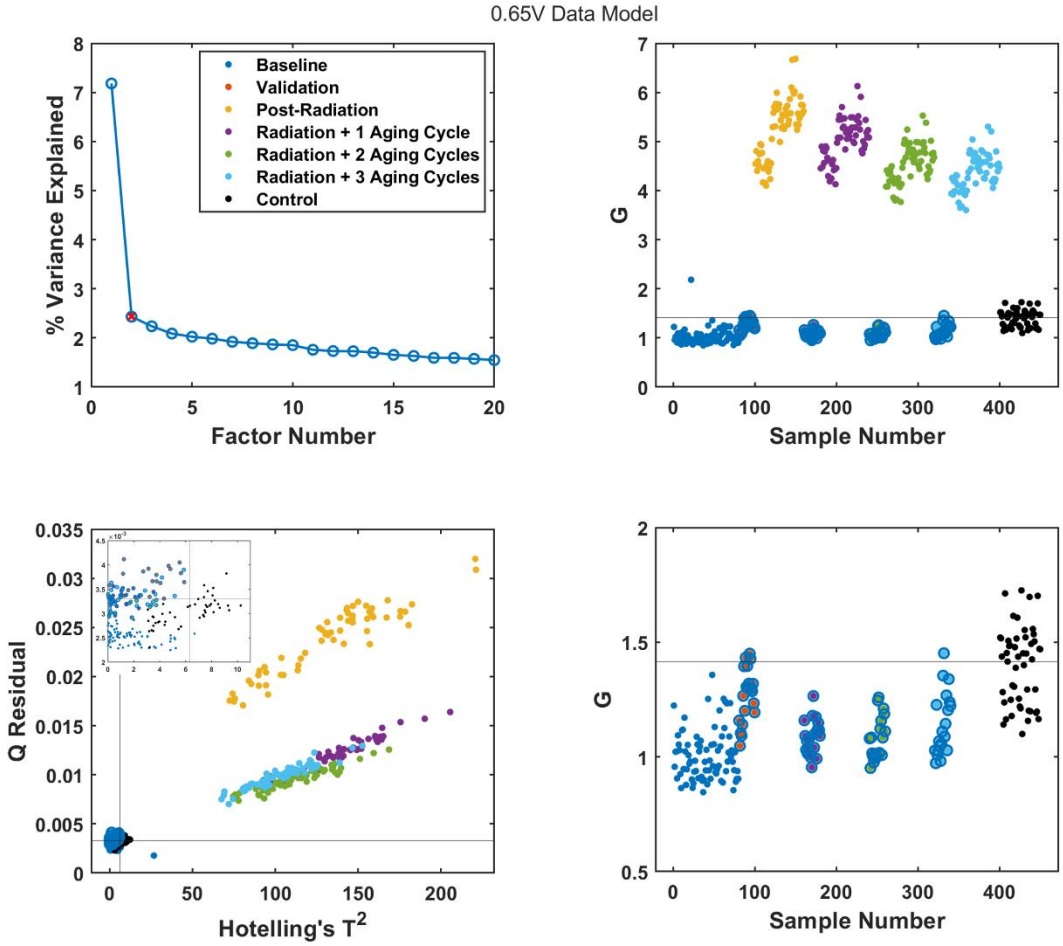


Figure 29. Classification model using 0.65V data (A) % Variance Explained with model rank marked (red "x") (B) G metric (C) Q-Residuals and Hotelling's T^2 (D) G metric, expanded view. Samples that have undergone aging but were *not* exposed to radiation are marked with a blue outline.

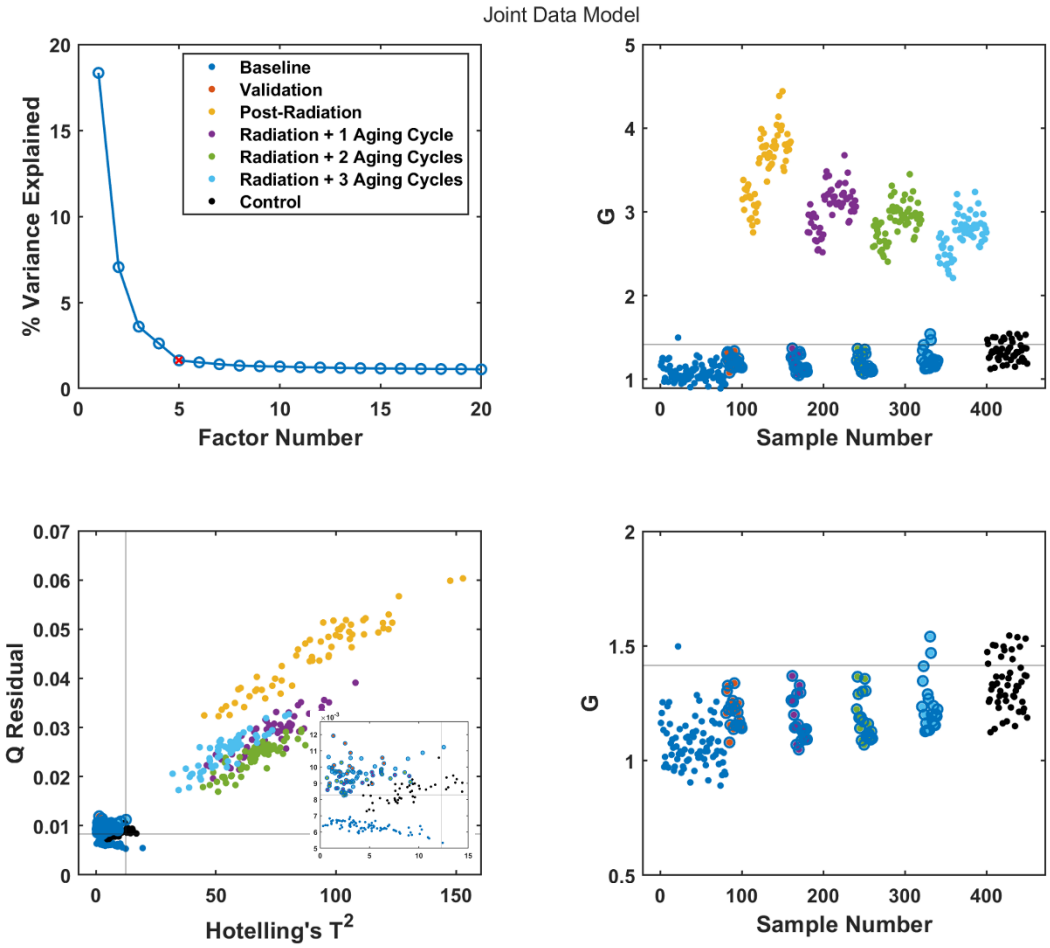


Figure 30. Classification model using 0.65, 1, and 3V data (A) % Variance Explained with model rank marked (red "x") (B) G metric (C) Q-Residuals and Hotelling's T^2 (D) G metric, expanded view. Samples that have undergone aging but were *not* exposed to radiation are marked with a blue outline.

4.3. SNL ASIC Aging

Similar to the analysis of the EC lot, data from several other SNL-manufactured ASIC lots were modeled to identify units that had been subjected to accelerated aging. These aged units are stand-ins for the anomalous units that would be identified during screening. Results of analysis for Lot 03 are discussed here.

Sample units from Lot 03 were aged for one to three cycles, and PSA data were collected after each aging cycle for comparison with data collected prior to aging. PC models were created for each biasing condition (0.65V, 1V, and 3V) as well as for different combinations of those voltages. The number of factors used in each model was identified using the % variance explained and cross-validation error (left column of Figure 31). Looking at the results in Figure 31, it is immediately apparent that the G metric is not appropriate for classification of these data. Despite seeing separation between aged and unaged data by the Q-Residual (center column), very few aged units are identified as "not pristine" by the G metric. As briefly mentioned above, the combination of Q-

Residuals and Hotelling's T^2 values into a single metric reduces the discriminatory power of the model in cases where only one metric appears to strongly indicate class membership.

In this case, focus can be turned to using the Q-Residual as the primary method of determining whether an unknown sample belongs in the "pristine" class. For these data, the models combining multiple bias conditions better separate aged from unaged data. Figure 32 shows the Q-Residuals and Hotelling's T^2 values for the model incorporating all three voltage conditions. This model visually shows a relatively clear separation between aged and unaged PSA data. However, the 95% confidence limit drawn for the Q-Residuals (marked in red) results in several false negatives (aged samples identified as unaged). If the limit is set to the lower 90% confidence limit (black line), many of these false negatives are avoided. This is seen in the confusion matrices shown in Table 8, where changing the confidence limit from 95% to 90% reduces false negatives from 24 samples to one sample.

In applications such as acceptance screening, high priority is placed on identifying any potential anomalous units. In exchange for this, it is often acceptable to risk higher levels of false positives (unaged units identified as aged) to reduce the rate of false negatives by lowering the detection limits (lowering the confidence limit in this case). These trade-offs are situationally dependent, and can guide the selection of decision boundaries for classification.

Using Q-Residuals alone might present a risk of missing some anomalies that present as more subtle changes in the device under study. To reduce that risk, the Hotelling's T^2 value, expected to detect those changes, can be incorporated into the model as a separate criterion. Rather than combining the Q-Residual and Hotelling's T^2 value in the same metric, devices can be assessed by the two metrics separately to ensure that anomalous units are likely to be detected.

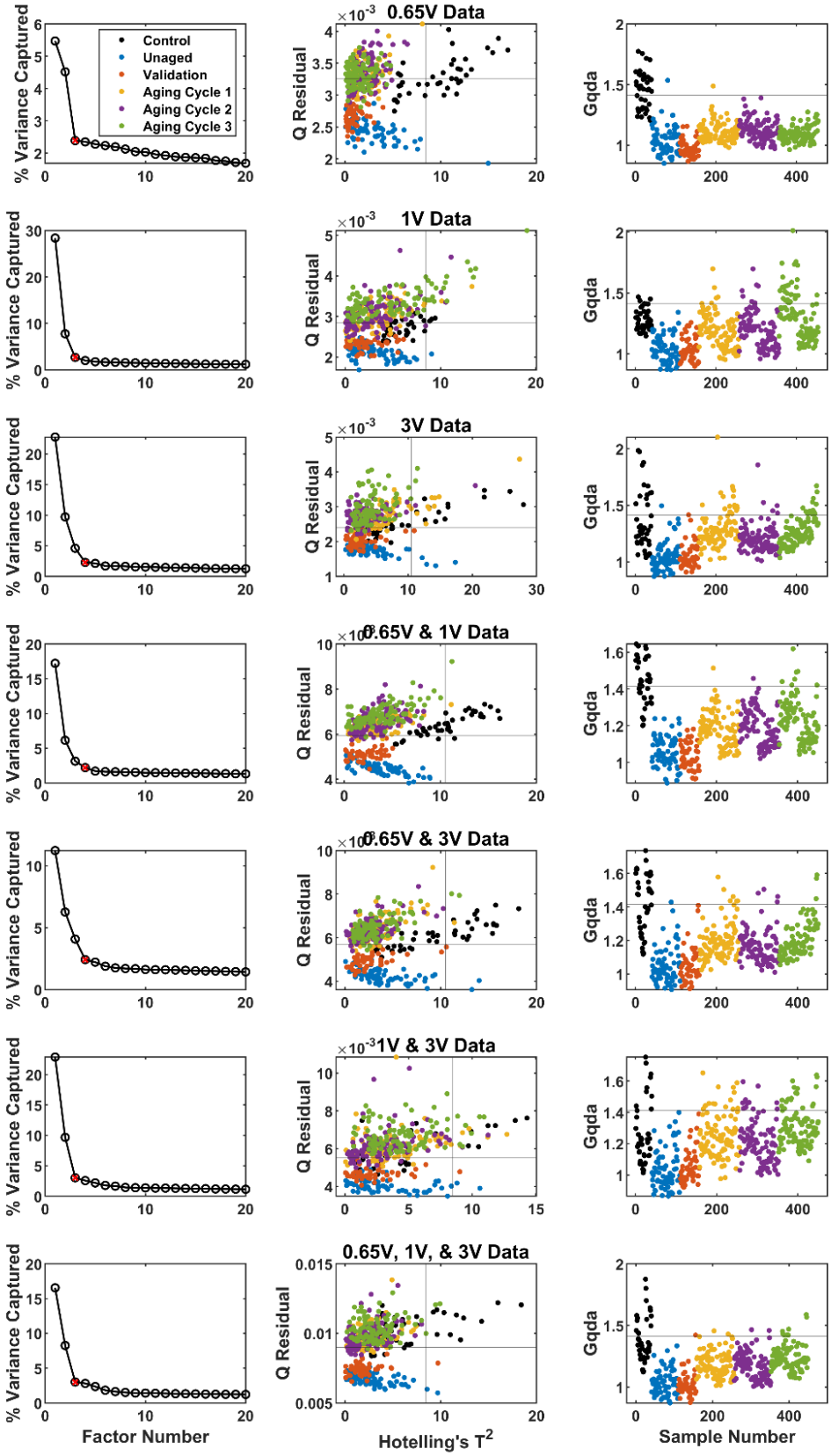


Figure 31. PC Models of Lot 03 ASIC Data

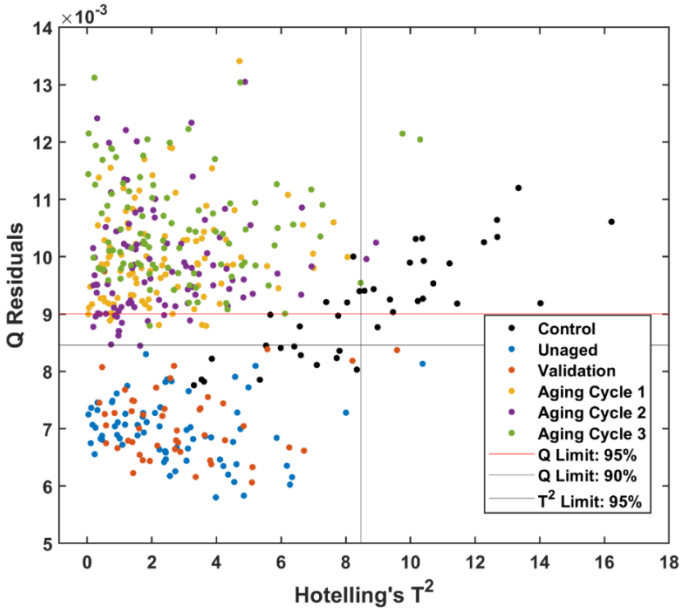


Figure 32. Q-Residuals and Hotelling's T^2 Values for 0.65, 1, and 3 V Model

Table 8. Confusion Matrices for ASIC Lot 03

95% Confidence Limit

	Predicted Pristine	Predicted Aged
Actual Pristine	**117**/117	**0**/117
Actual Aged	**24**/292	**268**/292

90% Confidence Limit

	Predicted Pristine	Predicted Aged
Actual Pristine	**117**/117	**0**/117
Actual Aged	**1**/292	**291**/292

4.4. Classification of Differential Line Drivers ESD Data

A classification model for AM26C31M Differential Line Drivers was developed following the methods laid out in the previous section. These data were previously investigated in the MDA report by P. Tangyunyong,[21] and the work here explores the utilization of an automated screening metric applied to the same data.

The data set consists of 100 AM26C31M Differential Line Drivers procured by MSWC Crane in two batches (85 units in batch 1, 15 units in batch 2). There are also 25 AM26C31IN and 25 AM26C31CN drivers procured from Digikey. The IN/CN model packaging is plastic, in contrast to the ceramic packaging of the M model. The PSA data were measured across several biasing conditions to interrogate different levels of the devices, identified here as Bias 1–4. The 85 AM26C1M Batch 1 devices were used to develop the baseline PCA model and detection thresholds. Prior to modeling, outlier data were removed. The same model was then used to screen the remaining PSA data for packaging differences or electrostatic discharge (ESD) insult.

All biasing conditions were easily able to identify AM26C31CN and AM26C31IN devices (Figure 33), with Bias 3 demonstrating the cleanest separation between the three groups. Data collected at Bias 1 show the greatest spread; this resulted in removing more outliers from the Bias 1 data, and a higher number of good units were rejected during the screening. PSA is well-suited to identifying these differences in the devices.

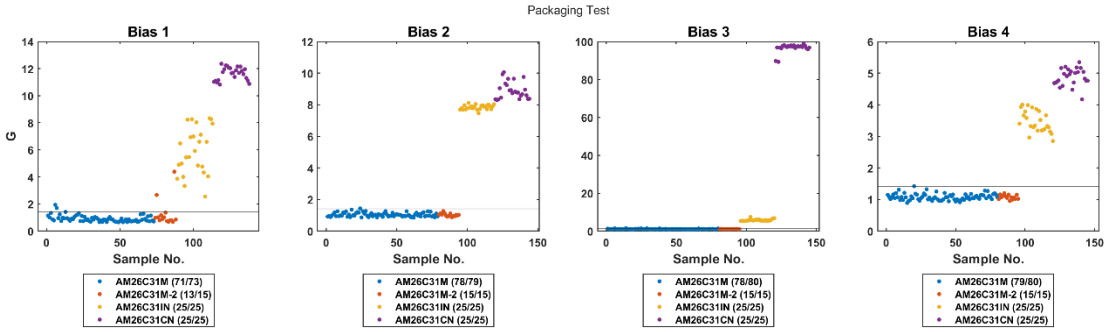


Figure 33. AM26C31M Model: Identification of AM26C31IN and AM26C31CN

Results are more varied for the devices exposed to ESD insults (Figure 34). The samples labeled "Validation" in Figure 34 were collected at the same time as the ESD samples but were not exposed to ESD, so they serve as controls here. For all four bias conditions, a few samples show consistently high values: 21, 22, and 23. As shown in Table 9, these units were subjected to the most extreme voltages, and all failed subsequent continuity testing. PSA was able to successfully identify these failed units at all bias conditions.

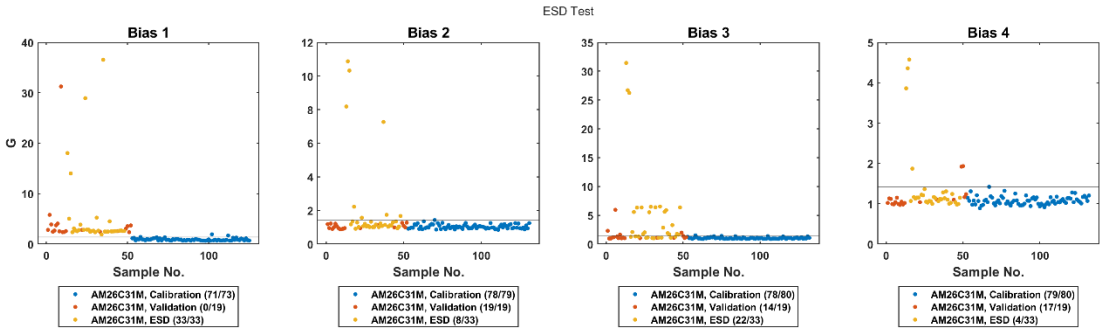


Figure 34. AM26C31M Model: Identification of ESD exposed samples

PSA data collected at Bias 1 showed a strong relationship to collection date, despite the centering performed to reduce the influence of instrument variability. This issue is readily apparent as both the validation and ESD-exposed samples scored above the detection threshold, resulting in a high rate of false positives. For that reason, results from the Bias 1 model are excluded from Table 9. PSA data collected at Bias 4 generally showed the least sensitivity to ESD exposure. Only the most extreme samples were identified as different from the baseline data.

PSA data collected under Bias 1, Bias 2, and Bias 3 show a set of intermediate data that score between the baseline values (or control values, in the case of Bias 1) and the extreme outliers noted above. This separation shows that some amount of ESD-induced damage can be identified in units prior to failure, and the nature of that identified damage might vary across biasing conditions, resulting in changes in effective sensitivity across the PSA data. In this case, it appears that Bias 3 is best suited for identifying the lower levels of ESD damage, correctly identifying 22 of 33 ESD-exposed units, and misidentifying 5 of 19 control samples. Though not all units exposed to ESD were identified, PSA appears sensitive enough to detect damage or device changes induced by exposure to ESD. Combination with multivariate analysis tools enables quick, automated screening of these changes.

PSA data from multiple biasing conditions (Bias 2–4) were also combined into a single model to explore if any increases in sensitivity could be achieved rather than using individual biasing conditions. Those results, shown in Figure 35, are overall quite similar to those for Bias 3, though

the visual spread in the data is smaller, suggesting that the combined model might be more robust to instrument variability.

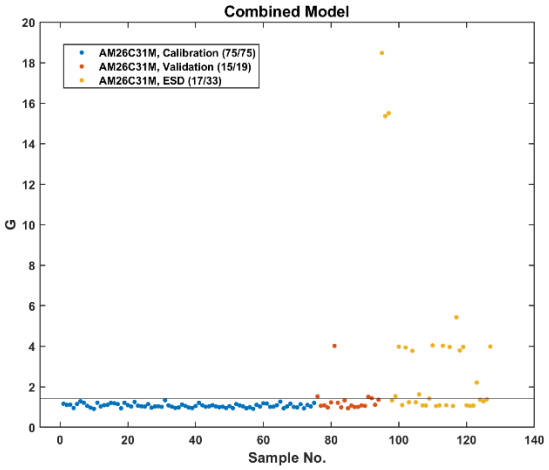


Figure 35. Classification of AM26C31M devices using multiple biasing voltages

Table 9. Classification Results for ESD Tested AM26C31 Devices

Sample	Voltage	Test Condition	Continuity after ESD	Bias 2, 3, or 4	Bias 2	Bias 3	Bias 4	Bias 2-4
21	4-8 KV	All Pins to GND	FAIL	Y	Y	Y	Y	Y
22	4-8 KV	All Pins to GND	FAIL	Y	Y	Y	Y	Y
23	4-8 KV	All Pins to GND	FAIL	Y	Y	Y	Y	Y
24	3500 V	1A to GND	PASS	N	N	N	N	N
25	4500 V	1A to GND	PASS	Y	N	Y	Y	Y
26	5000 V	1A to GND	PASS	Y	Y	Y	N	Y
27	5250 V	1A to GND	PASS	N	N	N	N	N
28	6000 V	1A to GND	PASS	Y	N	Y	N	Y
29	3500 V	1A to Vcc	PASS	N	N	N	N	N
30	4500 V	1A to Vcc	PASS	Y	Y	Y	N	Y
31	5000 V	1A to Vcc	PASS	Y	N	Y	N	N
32	5250 V	1A to Vcc	PASS	Y	N	Y	N	Y
33	6000 V	1A to Vcc	PASS	N	N	N	N	N
34	3500 V	1A to I/O	PASS	N	N	N	N	N
35	4500 V	1A to I/O	PASS	Y	N	Y	N	Y
36	5000 V	1A to I/O	PASS	Y	N	Y	N	Y
37	5250 V	1A to I/O	PASS	N	N	N	N	N
38	6000 V	1A to I/O	PASS	N	N	N	N	N
39	3500 V	1Y to GND	PASS	Y	N	Y	N	Y
40	4500 V	1Y to GND	PASS	N	N	N	N	N
41	5000 V	1Y to GND	PASS	Y	N	Y	N	Y
42	5250 V	1Y to GND	PASS	N	N	N	N	N
43	6000 V	1Y to GND	FAIL	Y	Y	Y	N	Y
44	3500 V	1Y to Vcc	PASS	Y	N	Y	N	Y
45	4500 V	1Y to Vcc	PASS	Y	Y	Y	N	Y
46	5000 V	1Y to Vcc	PASS	Y	N	Y	N	N
47	5250 V	1Y to Vcc	PASS	Y	N	Y	N	N
48	6000 V	1Y to Vcc	PASS	N	N	N	N	N
49	3500 V	1Y to I/O	PASS	Y	N	Y	N	Y
50	4500 V	1Y to I/O	PASS	N	N	N	N	N
51	5000V	1Y to I/O	PASS	Y	N	Y	N	N
52	5250 V	1Y to I/O	PASS	Y	N	Y	N	N
53	6000 V	1Y to I/O	PASS	Y	Y	Y	N	Y

“Y” denotes identified as anomalous by the classification model. “N” denotes identified as baseline.

This page left blank

5. PROPOSED WORKFLOWS FOR FUTURE APPLICATIONS

Based on the data investigated in this report, several recommendations can be made to develop and employ PSA models for both COTS electronics screening and surveillance evaluation. The general procedure for developing data models will be consistent across applications, though details will vary. This section lays out basic steps to building the models that could be used in a larger scale implementation of PSA. These activities will need to be implemented for every unique device, and use of the models across device lots and date codes will likely require model updating or recalibration to incorporate the additional variability, as PSA is sensitive to those changes.

5.1. Using PSA to Screen COTS Electronic Devices

The workflow to apply classification models to screening of COTS electronic devices is summarized in Figure 36.

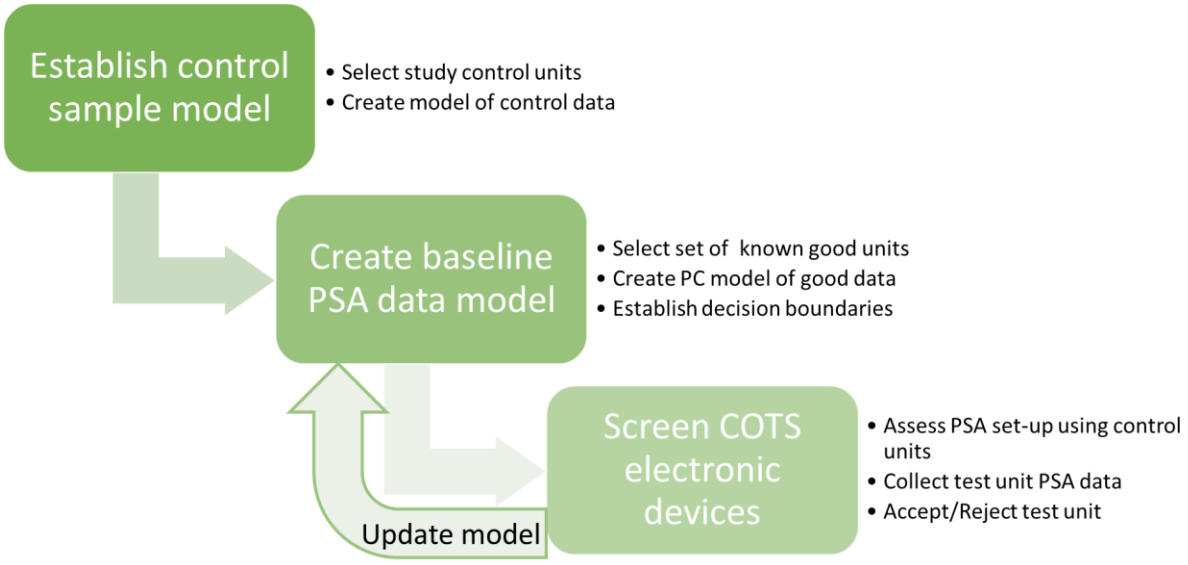


Figure 36. Screening Workflow

The detailed workflow is as follows:

1. **Select control samples**

   These samples will be used throughout data collection to monitor PSA instrument health and ensure that any changes in PSA signals are reflective of device change, rather than instrument set-up. The control samples should be similar to the devices under investigation, though they do not need to be identical (e.g., can be from another lot of similar devices). The samples should not be exposed to large environmental changes or other testing throughout the study duration to ensure that the assumption of unchanged status is not violated.

2. **Create control sample model**

   PSA data of the control samples collected at the beginning of the study can be used to form a control sample PCA model, similar to classification models of samples demonstrated in Section 4. At the start of new collection events, the control samples will be measured and compared to this initial model using the Q-Residuals and Hotelling's T^2 values to assess the PSA instrument set-up.

3. **Select baseline units to develop PC model**

   Baseline sample devices will be used to create the in-class boundaries using a PCA model of the device PSA data. These samples are preferably of the same lot as the target screening devices. If they are not of the same lot, further investigation will need to be undertaken to ensure that the model will be fully transferrable to PSA data of devices from a separate lot. All of the baseline samples should be known healthy units, while known bad units can be employed to test the developed model. PSA data collected from the baseline units should be collected over a few collection events to reduce the influence of changing experimental set-ups on the model. Any outlier PSA data of the baseline sample set should be identified and removed prior to PCA. The final PCA model of the PSA data can be used with the appropriate number of PCs to screen test samples.

4. **Establish decision boundaries**

   After the PC model of the baseline samples has been established, the decision boundaries can be determined. These boundaries will be used to make the determination of whether an unknown test article is acceptable or shows signs of anomalous behavior and should be rejected. The exact placement of these boundaries is based on acceptable levels of risk for false positives (rejecting good units) and false negatives (rejecting bad units). Generally, screening applications favor conservative thresholds. This means that the limits will ensure detection of any true negatives, with the tradeoff of rejecting some number of good parts. In the models demonstrated in this report, the threshold is determined as the confidence intervals for the Q-residuals and the Hotelling's T^2 values. When appropriate, these two metrics can be combined in a single metric: G.

   PSA data from any known anomalous units (accelerated aging, visually defective) can be used to test this metric. The decision boundaries can be adjusted based on the results of that test.

5. **Employ model**

   At the beginning of each collection event, PSA data of the control samples should be collected and compared against the control model. If the control samples are not consistent with the model, adjustments to the PSA experimental set-up can be made to reduce changes in the PSA data. If the changes in the control model persist, a correction factor can be determined for the data in this collection based on the control data. Performing this data quality check will prevent collection of uninformative sample PSA data and improve model performance.

   Sample PSA data will then be compared to the baseline model using the decision boundaries established in Step 4. Units will be accepted or rejected based on the PSA data model.

6. **Update models**

   If models are used for many more units, additional known good parts can be incorporated to increase the variability included in the initial model. This will make the model more robust in its continued application.

5.2. Using PSA to Monitor Aging of Electronic Devices

A summary workflow for monitoring aging of test devices is shown in Figure 37, and detailed below. This workflow is similar to that employed in screening applications, but incorporate the different modeling approach with the aim of estimating device age.

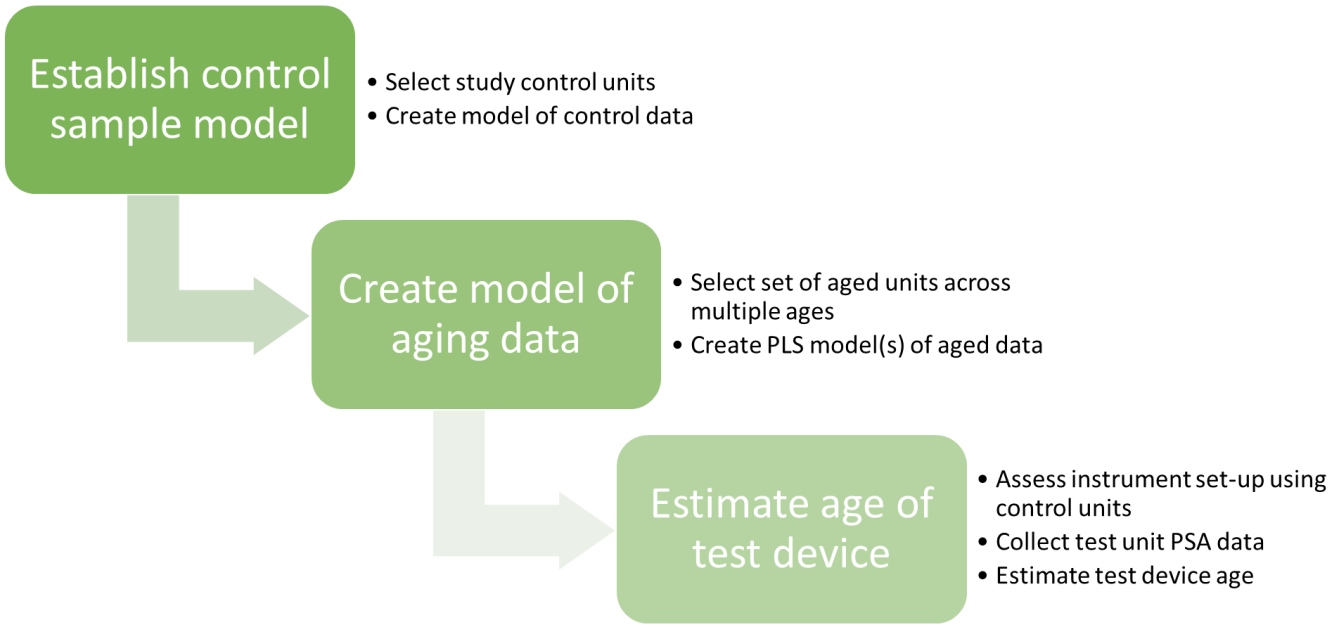


Figure 37. Aging Workflow

1. **Select control samples**

   These samples will be used throughout data collection to monitor PSA instrument health and ensure that any changes in PSA signals are reflective of device change, rather than instrument set-up. The control samples should be similar to the devices under investigation, though they do not need to be identical (e.g., can be from another lot of similar devices). The samples should not be exposed to large environmental changes or other testing throughout the study duration to ensure that the assumption of unchanged status is not violated.

2. **Create control sample model**

   PSA data of the control samples collected at the beginning of the study can be used to form a control sample PCA model, similar to classification models of samples demonstrated in Section 4. At the start of new collection events, the control samples will be measured and compared to this initial model using the Q-Residuals and Hotelling's T^2 values to assess the PSA instrument set-up.

3. **Select baseline units to develop PLSR model**

   Units selected to develop the aging model should be representative of the units that will be subjected to testing in the future. These units need to span a range of accelerated aging times (or real-time aging) expected to induce signs of aging in the devices. Outlier PSA data should be removed prior to model calibration. The PLSR model can then be calibrated following the method in Section 3.4. The PLSR model should be used to predict ages of control samples; these samples should be consistent across time to demonstrate that the model is predictive of age, rather than sample collection date.

4. **Estimate device ages**

   At the beginning of each collection event, PSA data of the control samples should be collected and evaluated in the PLSR model. If the control samples are not consistent with earlier control data, adjustments to the PSA experimental set-up can be made to reduce changes in the PSA data. If the changes in the control model persist, a correction factor can be determined for the

data in this collection based on the control data. Performing this data quality check will prevent collection of uninformative sample PSA data and improve model performance.

Sample PSA data will then be evaluated by the PLSR model. Estimated sample age can be compared to the known sample age to evaluate whether the unit is aging as expected.

6. CONCLUSIONS

This report investigated existing PSA data sets to understand the utility of PSA combined with multivariate data analysis to evaluate COTS electronics components. Section 2 evaluated PSA data collected from a nominally unchanged set of control samples over four years. Analysis of these data demonstrated how changes in experimental set-up had the potential to influence future model performance, and showed that a handful of sample units were needed to establish baseline predictive models.

The application of PSA to evaluate COTS device age was explored in Section 3. These data demonstrated how changes in experimental set up could influence models, requiring correction for collection event or the need to evaluate experimental set-ups using control samples prior to data collection. While the PSA data of Zener diodes showed substantial spread in the PLSR models, aging could still be modeled to predict age of the validation samples. Aging models could be used in the future to evaluate real-time aging data to assess the validity of accelerated aging data, and to understand whether samples are effectively the age they are expected to be. Discrepancies between predicted age and known sample ages might point to unexpected aging behavior in samples, and could drive further work to fill in any knowledge gaps. A technical advance for the method of developing the algorithms demonstrated in this section was filed (SD# 16263).

The application of PSA to screen COTS electronics was explored in Section 4. PCA models combined with a classification metric based on residual model error (Q-residuals) was able to successfully detect sample units exposed to radiation or accelerated aging conditions. The simplicity of implementing these models to make acceptance decisions puts PSA in position for use as a 100% screening tool for COTS electronic components.

The lessons learned from these example implications formed the basis for workflows laid out in Section 5. Future development of these workflows could target combined use of the screening and aging models. First, units could be screened for defects or other gross anomalies. Following that, aging models sensitive to smaller changes relating to sample age could be employed to determine remaining expected lifetime of these components. Both applications have the potential to enable reliable use of COTS electronic devices throughout the stockpile, using a fast, non-destructive testing device partnered with multivariate data analysis.

This page left blank

REFERENCES

1. Tangyunyong, P.C.J., E. I.; G. M. Loubriel; J. Beutler; D. M. Udoni; B. S. Paskaleva; T. E. Buchheit *Power Spectrum Analysis (PSA)*. in *43rd International Symposium for Testing and Failure Analysis*. 2017.
2. Tangyunyong, P.C.J., E. I.;Stein, D. J., *Power Spectrum Analysis for Defect Screening in Integrated Circuit Devices*. 2015: United States.
3. Dong, L.L., Gavin R.; Duncan, John C.; Brereton, Richard G., *Disjoint hard models for classification*. Journal of Chemometrics, 2010(24): p. 273-287.
4. van der Maaten, L.H., Geoffrey, *Visualizing Data using t-SNE*. Journal of Machine Learning Research, 2008. **9**: p. 2579-2605.
5. Brereton, R.G., *Chemometrics for Pattern Recognition*. 2009, New York, UNITED KINGDOM: John Wiley & Sons, Incorporated.
6. Brereton, R.G., *One-class classifiers*. Journal of Chemometrics, 2011(25): p. 225-246.
7. Miller, P.S., Ronald E.; Heckler, Charles E., *Contriubtion Plots: A Missing Link in Multivariate Quality Control*. Applied Math and Computer Science, 1998. **8**(4): p. 775-792.
8. Wold, S.S., Michael; Eriksoon, Lennart, *PLS-regression: a basic tool of chemometrics*. Chemometrics and Intelligent Laboratory Systems, 2001. **58**: p. 109-130.
9. Gemperline, P., *Practical Guide to Chemometrics*. Second ed. 2006: Taylor & Francis Group.
10. Brereton, R.G.L., Gavin R., *Partial least squares discriminant analysis: taking the magic away*. Journal of Chemometrics, 2014. **28**: p. 213-225.
11. Esbensen, K.H., *Multivariate Data Analysis-In Practice*. Fifth ed. 1994, Oslo, Norway: Camo.
12. Multari, R.C., C.; Ferrizz, R.M.; Patel, N.; Cummings, P.G.; Bujewski, E.G.; Walla, L.A.; Dailey, M.; Howard, J.S.; Nelson, M.; Martin, S.; Miller, L.M.; Marklin, S.; Kalinich, D.A.; Phillips, H.C.; Ray, J, *Automated Algorithms for Predicting Trends and Identifying Subpopulations in Neutron Generator (NG) Production*. 2021.
13. Multari, R.A.C., D. A.; Dupre, J. M.; Gustafson, J. E., *The use of Laser-induced Breakdown Spectroscopy (LIBS) for distinguishing between bacterial pathogen species and strains*. Applied Spectroscopy, 2010(64): p. 750-759.
14. Multari, R.A.C., D. A., *Methods for Forming Recognition Algorithms for Laser-Induced Breakdown Spectroscopy*. 2011: United States.
15. Multari, R.A.C., D. A.; Bostian, M. L., *Use of Laser-induced Breakdown Spectroscopy for the differentiation of pathogens and viruses on substrates*. Applied Optics, 2012(51): p. B57-B64.
16. Multari, R.A.C., D. A.; Scott, T.; Kendrick, P., *Detection of pesticides and dioxins in tissue fats and rendering oils using Laser-induced Breakdown spectroscopy (LIBS)*. Journal of Agriculture and Food Chemistry, 2013(61): p. 2348-2357.
17. Multari, R.A.C., D. A.; Dupre, J. M.; Gustafson, J. E., *The detection of biological contaminants on foods and food surfaces using Laser-induced Breakdown Spectroscopy (LIBS)*. Journal of Agriculture and Food Chemistry, 2013(61): p. 8687-8694.
18. Multari, R.C., D. A.; Duncan, R.; Young, S. *Laser-induced Breakdown Spectroscopy as a rapid, in-situ clinical diagnostic*. in *21st International conference on biodetection technologies, 2013: technological advances in detection and identification of biological threats*. 2013. Alexandria, VA.
19. Multari, R.A.C., D. A.; Bostian, M. L.; Dupre, J. M.; Gustafson, J. E., *Proof-of-principle for a real-time pathogen isolation media diagnostic: The use of Laser-induced Breakdown Spectroscopy (LIBS) to discriminate bacterial pathogens and antimicrobial-resistant Staphylococcus aureus strains grown on blood agar*. Journal of Pathogens, 2013(11).

20. Multari, R.A.C., D. A., *Laser-induced Breakdown Spectroscopy (LIBS) as a potential tool for food safety*, in *Food poisoning: Outbreaks, bacterial sources and adverse health effects*, P.C. Ray, Editor. 2015, Nova Publishers: New York, NY. p. 63–74.
21. Tangyunyong, P., *Feasibility of Using PSA for Counterfeit Detection of Specific Integrated Circuits*.

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Charles Richardson	0491	cbricha@sandia.gov
Coby Davis	1854	cldavis@sandia.gov
Yibin Zhang	2434	yibzhan@sandia.gov
Justin Serrano	2491	jrserra@sandia.gov
Anthony Colombo	5212	apcolom@sandia.gov
Greg White	7254	gvwhite@sandia.gov
Juan Ubiera	7257	jrubier@sandia.gov
Sean Hamilton	7555	srhamil@sandia.gov
Technical Library	1911	sanddocs@sandia.gov

This page left blank

This page left blank

Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.